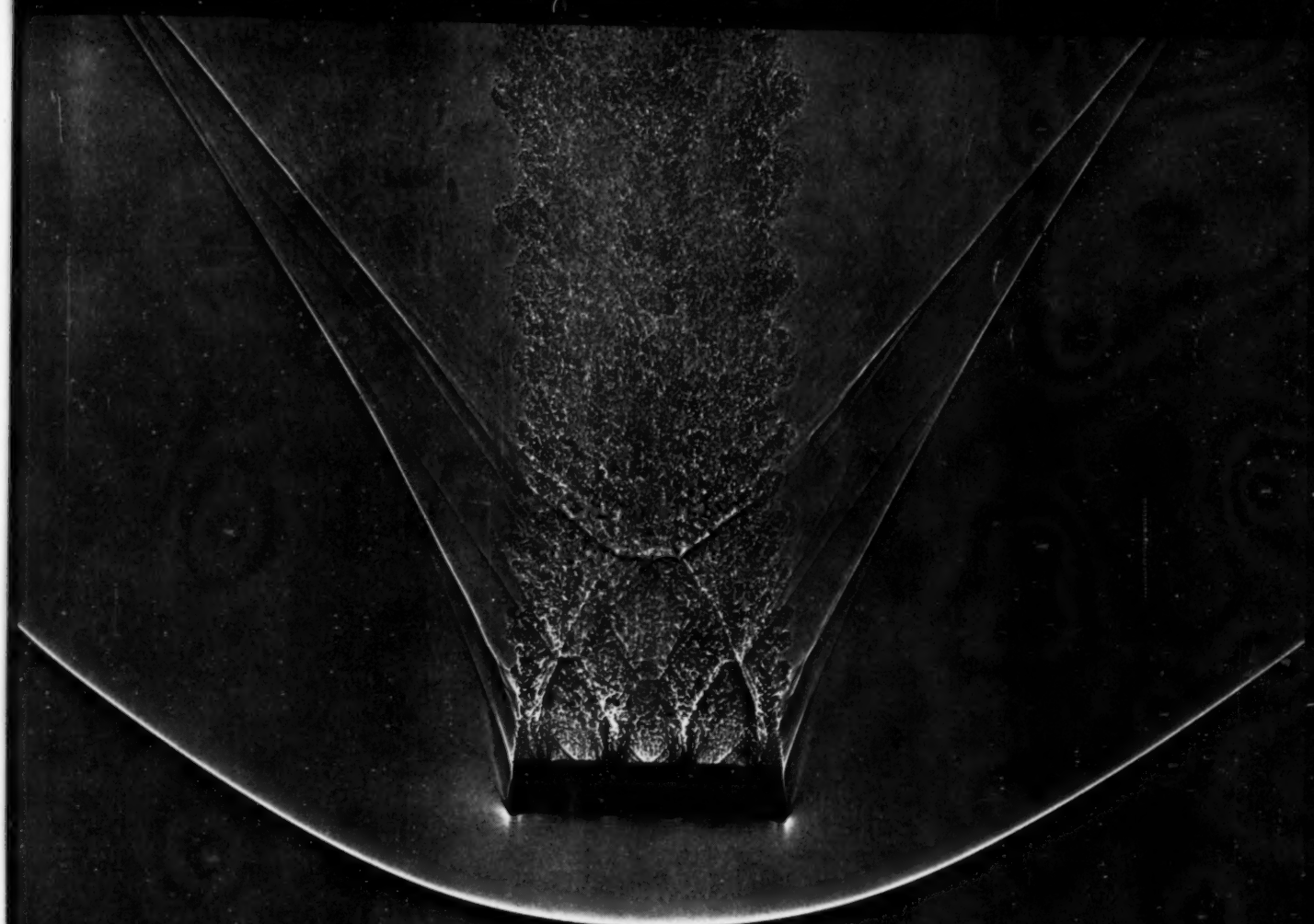
JOURNAL OF APPLIED PHYSICS



.... new X-ray instrumentation for
INDUSTRIAL PROCESSING CONTROL

Norelco AUTROMETER

The most recent addition to the research and control instruments developed by North American Philips Company, is the Autrometer—an automatic multi-element indexing spectrometer for process control industries. It will provide process industries with an accurate device capable of determing the presence of as many as twelve elements in a sample material, and may be set for any element from a range of over seventy-one elements in the periodic table.

The Autrometer is indispensable in the precision control of quality. One important advantage of the new instrument is that it will provide accurate analysis of products in any critical stage during the course of manufacture. These include: incoming raw materials; constituents in various stages of combination; also the finished product. As such, the Autrometer gives rapid indication of the necessity for the addition or removal of elements to guarantee uniformity of mixes, batches or the degree of purity of any of the components. Thus, fullest maintenance of product standards and product specifications can be fully controlled to assure the industrialist predeterminable, repetitive function and performance of the end product.

Facility of use, speed of operation and rapidity of analysis inherent in the Autrometer, brings to industry the first rapid-production automatic processing analyzer.

A Norelco PRODUCT NORTH AMERICAN PHILIPS COMPANY, INC.

Serving Science and Industry

750 South Fulton Avenue, Mount Vernon, N. Y.

In Canada: Rogers Majestic Electronics Ltd., 11-19 Brentcliffe Road, Leaside, Toronto 17, Ontario

Journal
of
Applied Physics

Volume 26, Number 1 January, 1955

Editorial

Under the pressure of duties ever more abundant and important, Dr. Elmer Hutchisson, founding Editor of this Journal, has had, with this issue, to withdraw his name from the masthead. Everyone will regret this. As far as the American Institute of Physics is concerned, he could have stayed on indefinitely.

When the Institute was founded in 1931, it was a fact recognized by leading physicists that their science and their colleagues were but poorly recognized in industrial circles. There were notable exceptions to this, of course. Certain companies such as the Bell Telephone Laboratories, the General Electric Company, The Gulf Research and Development Corporation, and a very few others had taken on physicists in great numbers, and the record shows that they made no mistake. By and large, however, American industry employed physicists in very small proportion to the actual dependence of industry on this basic science.

Accordingly, one of the purposes of the Institute from the beginning was to promote "physics in industry." The *Journal of Applied Physics* was established to do just that. This is where Dr. Hutchisson came in. At the Institute's request, he took leave of absence from his post on the physics faculty of the University of Pittsburgh, came to the AIP office in New York as Assistant Director of the Institute, and spent about a year planning, promoting, and finally editing this Journal. After the first few issues had been published early in 1937, he returned to Pittsburgh as Acting Head and, later, Head of the Department of Physics.

During Dr. Hutchisson's tenure of office as Assistant Director, he largely planned and managed the Fifth Anniversary meeting of the AIP, held in New York in October 1936. The meeting was built around the same theme as was later to guide the Journal—presentation of the achievements of physics in industry and recognition of the resourcefulness of physics in industrial research and development. This meeting—a joint one of all of the Member Societies—was the springboard of the Journal. Dr. Hutchisson's leading part in the meeting was a portent of his breadth of interest. He was not "just the Editor of JAP." He was the protagonist of the concept behind the Journal. He saw the Journal originally not as a repository of research reports, but as a means to further a constructive cause.

Although, as he said editorially last July, "During the years of World War II . . . the Journal became more of an archive publication," this was really the result of success in attaining the original objective. Physics now has its role in industry; 36 percent of physicists are employed by private industry, and the outpouring of their researches has flooded the pages available to record them.

Dr. Hutchisson continued as Editor when in 1945 he went to his Alma Mater, Case Institute of Technology, as Dean of the Faculty. He continued during the anxious and strenuous war years when, as Technical Aide to the National Defense Research Committee, he played a leading part in the highly important undersea warfare project. He even continued when for a later time he was Acting President of Case. It was not until then that he required assistance and obtained it through the appointment of Professor Earle C. Gregg as Assistant Editor in 1950.

The advantages to the civilized community, to American industry, and to physics which have accrued from the leadership of Dr. Hutchisson are great and beyond our present power to estimate. This editorial expression of recognition in the name of the American Institute of Physics is placed here in the record with the hearty approval of the new Editor, Robert L. Sproull.

HENRY A. BARTON
Director of the Institute

Čerenkov Radiation from Extended Electron Beams*

M. Danos†
Columbia Radiation Laboratory, Columbia University, New York 27, New York
(Received July 8, 1954)

The Čerenkov radiation emitted by a bunched electron beam passing along a dielectric material at a close distance is calculated for a flat beam near a plane surface, a flat beam between two plane surfaces and for a circularly cylindrical beam within a cylindrical hole.

I. INTRODUCTION

THE nature of the Čerenkov radiation has been exhaustively investigated in a number of papers.[1–5] All these investigations dealt with the radiation emitted by a single electron and with the properties of the radiation process. The present paper is mainly concerned with the case in which radiation is produced by an electron beam rather than by a single electron. The electron beam is treated as consisting of a continuous charge density moving with a constant velocity v common to all parts of the beam. That means that the distribution of the charge density is stationary in the system in which the beam is at rest, i.e. the beam moves as a rigid body. It will be assumed that the beam is bunched; therefore the charge density is a periodic function of space and time. Then the calculation yields directly the coherent radiation connected with the bunching of the beam. The incoherent part of the radiation can be obtained by assuming the beam to be modulated with a suitable random noise spectrum.

II. CALCULATIONS

The calculations proceed in the following manner. First a solution of the inhomogeneous wave equation is obtained with fields vanishing at infinity. This represents the field associated with the electron beam in the absence of the dielectric. Then a sufficiently general solution of the homogeneous wave equation is found for both the vacuum space, i.e. the space through which the electron beam moves, and the dielectric, with appropriate behavior at infinity. Finally the fields are matched at the dielectric-vacuum interface. The fields so obtained represent the solution of the problem. From these fields all quantities of interest can be calculated.

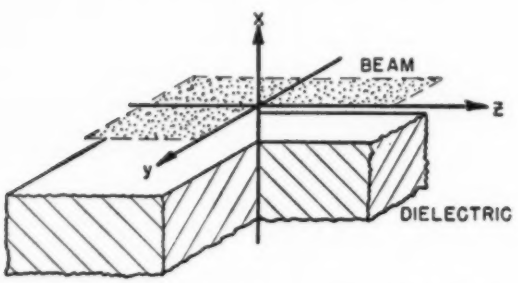


Fig. 1. Flat electron beam moving over dielectric.

A. Plane Geometry

The simplest case, which is probably of the most importance for practical application, is that in which a flat electron beam moves over a dielectric at a short distance. (See Fig. 1.) In this case the charge density is given by

$$\rho=\rho_0\delta(x)[1+\alpha\cos(kz-\omega t)] \tag{1}$$

with

$$k=\omega/v \tag{2}$$

where $v=c\beta$ is the velocity of the electrons and α is the Fourier-coefficient of the charge density corresponding to the frequency ω. The charge density has been written in this form in order to relate it to the average dc charge density represented by ρ_0 in Eq. (1). The beam is assumed to be infinitesimally thin.

The potentials in free space, i.e., for $x>-d$, obey the field equations

$$\Box V\equiv\frac{\partial^2V}{\partial x^2}+\frac{\partial^2V}{\partial y^2}+\frac{\partial^2V}{\partial z^2}-\frac{1}{c^2}\frac{\partial^2V}{\partial t^2}=-4\pi\rho \tag{3}$$

$$\Box\mathbf{A}=-4\pi\rho\frac{\mathbf{v}}{c}. \tag{4}$$

In our case $\mathbf{v}=v\mathbf{z}_0=$const ($\mathbf{z}_0$ unit vector). Therefore,

$$\mathbf{A}=\beta V\mathbf{z}_0 \tag{5}$$

holds and it is sufficient to obtain a solution of Eq. (3), i.e. to determine the scalar potential. The vector potential then can be calculated using Eq. (5).

The solution of Eqs. (3) and (4) using the ac part of the charge density only, is given by

$$V=(\Phi/q)\cos(kz-\omega t)e^{-q|x|} \tag{6}$$

$$\mathbf{A}=(\Phi\beta/q)\cos(kz-\omega t)e^{-q|x|}\mathbf{z}_0. \tag{7}$$

* Research supported jointly by the Signal Corps, Office of Naval Research, and the U. S. Air Force.
† Present address: National Bureau of Standards, Betatron Section, Washington, D. C.
[1] P. Čerenkov, Compt. rend. Acad. Sci. USSR 2, 457 (1934).
[2] I. Frank and Ig. Tamm, Compt. rend. Acad. Sci. USSR 14, 109 (1937).
[3] Ig. Tamm, J. Phys. Acad. Sci. USSR 1, 439 (1939).
[4] V. L. Ginsberg and I. M. Frank, Doklady Acad. Nauk USSR 56, 699 (1947).
[5] G. Beck, Phys. Rev. 56, 699 (1947).

Here,

$$q^2 \equiv k^2\vartheta^2 = k^2(1-\beta^2) \tag{8}$$

$$\Phi = 2\pi\alpha\rho_0. \tag{9}$$

Equations (6) and (7) can be easily verified using the well known identity:

$$\frac{d^2|x|}{dx^2} = 2\delta(x).$$

For the solution of the homogeneous equation

$$\Box \mathbf{A}_1 = 0 \tag{10}$$

we make the trial solution

$$\mathbf{A}_1 = [A_x \mathbf{x}_0 \cos(kz-\omega t+\varphi_x) + A_1 \mathbf{z}_0 \cos(kz-\omega t+\varphi)]e^{-qx} \tag{11}$$

$$V_1 = 0.$$

Inserting Eq. (11) in Eq. (10) we first obtain Eq. (8). From the Lorentz condition, $\mathrm{div}\mathbf{A}_1 = 0$, follows $A_x = A_1/\vartheta$; $\varphi_x = \varphi + \pi/2$, and so

$$\mathbf{A} = (A_1/\vartheta)e^{-qx}[\vartheta \mathbf{z}_0 \cos(kz-\omega t+\varphi) - \mathbf{x}_0 \sin(kz-\omega t+\varphi)]. \tag{12}$$

The vacuum fields associated with the electron beam, calculated from Eqs. (6) and (7) are

$$\mathbf{E}_0 = \Phi e^{-q|x|}[\vartheta \mathbf{z}_0 \sin(kz-\omega t) + \mathrm{sign}x \mathbf{x}_0 \cos(kz-\omega t)] \tag{13}$$

$$\mathbf{H}_0 = \mathrm{sign}x \Phi\beta e^{-q|x|}\mathbf{y}_0 \cos(kz-\omega t) \tag{14}$$

while Eq. (12) yields

$$\mathbf{E}_1 = -(A_1/\vartheta)\kappa e^{-qx}[\vartheta \mathbf{z}_0 \sin(kz-\omega t+\varphi) + \mathbf{x}_0 \cos(kz-\omega t+\varphi)] \tag{15}$$

$$\mathbf{H}_1 = -(A_1/\vartheta)\kappa\beta e^{-qx}\mathbf{y}_0 \cos(kz-\omega t+\varphi). \tag{16}$$

The function sign x is defined by

$$\mathrm{sign}\, x \equiv \begin{cases} +1 & \text{for } x>0 \\ 0 & \text{for } x=0 \\ -1 & \text{for } x<0 \end{cases}$$

and the abbreviation

$$\kappa \equiv \omega/c = 2\pi/\lambda = k\beta \tag{17}$$

has been used. In Eqs. (15) and (16) the quantities A_1 and φ are adjustable parameters which will be used to match the fields at the boundary.

The vector potential inside of the dielectric must obey the equation

$$\Box_D \mathbf{B} \equiv \frac{\partial^2\mathbf{B}}{\partial x^2} + \frac{\partial^2\mathbf{B}}{\partial y^2} + \frac{\partial^2\mathbf{B}}{\partial z^2} - \frac{\epsilon}{c^2}\frac{\partial^2\mathbf{B}}{\partial t^2} = 0. \tag{18}$$

We assume for B both an incoming and an outgoing wave. A solution obeying both the Lorentz condition and (18) is given by

$$\mathbf{B} = B_1\left(\mathbf{z}_0 - \frac{1}{\tau}\mathbf{x}_0\right) \times \cos[kz + p(x+d) - \omega t + \psi_1] + B_2\left(\mathbf{z}_0 + \frac{1}{\tau}\mathbf{x}_0\right)\cos[kz - p(x+d) - \omega t + \psi_2] \tag{19}$$

where

$$p^2 \equiv k^2\tau^2 = k^2(\epsilon\beta^2 - 1). \tag{20}$$

The first term in Eq. (19) represents an incoming wave. Therefore both B_1 and ψ_1 are arbitrary. The second term represents an outgoing wave with B_2 and ψ_2 to be determined by the matching of the fields. The fields calculated from Eq. (19) are

$$\mathbf{E}_D = -\kappa B_1\left(\mathbf{z}_0 - \frac{1}{\tau}\mathbf{x}_0\right)\sin[kz + p(x+d) - \omega t + \psi_1] - \kappa B_2\left(\mathbf{z}_0 + \frac{1}{\tau}\mathbf{x}_0\right)\sin[kz - p(x+d) - \omega t + \psi_2] \tag{21}$$

$$\mathbf{H}_D = \frac{\epsilon\kappa\beta}{\tau}B_1\mathbf{y}_0 \sin[kz + p(x+d) - \omega t + \psi_1] - \frac{\epsilon\kappa\beta}{\tau}B_2\mathbf{y}_0 \sin[kz - p(x+d) - \omega t + \psi_2]. \tag{22}$$

Inserting the fields as given by Eqs. (13), (14), (15), (16), (21), and (22) into the matching conditions: at the boundary, i.e. at $x = -d$

$$E_{\text{tang vac}} = E_{\text{tang diel}}; \quad H_{\text{tang vac}} = H_{\text{tang diel}} \tag{23}$$

a set of four simultaneous equations for the unknown quantities is obtained. One finds for the outgoing wave

$$B_2^2 = \frac{2\Phi^2\tau}{\epsilon\kappa^2}e^{-2qd}\sin\zeta + B_1^2 + \frac{2\Phi B_1}{\epsilon\kappa}e^{-qd}\sin\xi[\epsilon\vartheta \sin\psi_1 + \tau\cos\psi_1] \tag{24}$$

$$\sin\psi_2 = \frac{\Phi\vartheta}{B_2\kappa}e^{-qd}\sin\zeta + \frac{B_1}{B_2}\sin(\zeta+\psi_1) \tag{25}$$

$$\cos\psi_2 = \frac{-\Phi\vartheta}{B_2\kappa D}e^{-qd}\sin\zeta + \frac{B_1}{B_2}\cos(\zeta+\psi_1). \tag{26}$$

The phase shift ζ is given by

$$\sin\zeta = \frac{2D}{D^2+1}; \quad \cos\zeta = \frac{D^2-1}{D^2+1} \tag{27}$$

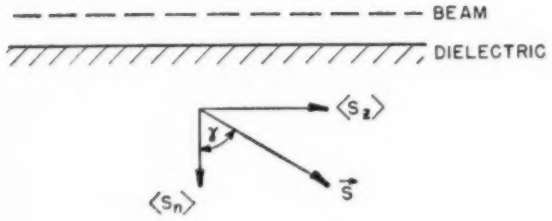


FIG. 2. Components of the Poynting vector.

where

$$D=\frac{\epsilon\vartheta}{\tau}. \tag{28}$$

The energy radiated by the electron beam is given by the time average of the Poynting vector:

$$\langle \mathbf{s} \rangle = \frac{c}{8\pi}\frac{\kappa^2\epsilon\beta}{\tau}B_2^2\left[-\mathbf{x}_0+\frac{1}{\tau}\mathbf{z}_0\right]. \tag{29}$$

The z component of the Poynting vector describes an energy flux which remains in the vicinity of the electron beam and is not radiated away. This part remains finite if the velocity of the electrons decreases below the value given by the Čerenkov condition $\tau^2>0$. The energy radiated away is given by the component of the Poynting vector normal to the beam. (See Fig. 2.)

The case of no incoming radiation is of special interest. Here the normal component of the Poynting vector can be written in the form

$$\langle s_n \rangle = 2\pi c\alpha^2\rho_0^2\beta^2 e^{-2qd}\eta(\epsilon,\beta) \tag{30}$$

with

$$\eta(\epsilon,\beta)=\frac{\epsilon}{\beta}\frac{\vartheta^2\tau}{\epsilon\vartheta^2+\tau^2}. \tag{31}$$

Equation (31) shows that $\langle s_n \rangle$ vanishes for $\beta\to 1$. This is due to the Lorentz contraction of the primary field which manifests itself through the factor ϑ in the z component of (13).

For practical purposes it is convenient to express the charge density through the current. Let $N\approx 6.67\times 10^{18}$ be the number of electrons in one coulomb, and let

$$R_0=\frac{2\pi N^2 e^2}{c}\times 10^{-7}\approx 189\Omega. \tag{32}$$

Then we obtain finally

$$\langle s_n \rangle = R_0(I\alpha)^2 e^{-2qd}\eta(\epsilon,\beta)\frac{\text{watts}}{\text{cm}^2}\text{ for } I \text{ in }\frac{\text{amp}}{\text{cm}}. \tag{33}$$

The dc current in the beam divided by the width of the beam is I. The quantity $R_0\eta(\epsilon\beta)$ can be interpreted as an intrinsic radiative resistance for Čerenkov radiation in this geometry.

Let us consider an example. We assume: $\epsilon=100$; $\beta=0.2$ (10 kev electrons); $d=3\times 10^{-2}$ cm; $\lambda=1$ cm; $I=1$ mA/cm. If the bunching is complete, i.e., if the bunches are short compared to $1/k$ then $\alpha=2$. With these values we obtain

$$2qd=\frac{4\pi}{\beta}\frac{d}{\lambda}(1-\beta^2)^{\frac{1}{2}}\approx 61.6\frac{d}{\lambda}=1.85 \tag{34}$$

and

$$\langle s_n \rangle = 11\times 10^{-6} W/\text{cm}^2.$$

This is to be compared with the kinetic energy in the considered beam which is $10W$/cm.

The angle γ at which the radiation is emitted (Fig. 2) is given by

$$\tan\gamma=\frac{\langle s_z \rangle}{\langle s_n \rangle}=\tau. \tag{35}$$

In our example therefore $\gamma=60°$. The quantity D in Eq. (28) here has the value $D=56$.

One of the assumptions used in the derivation was that the charge is concentrated in a sheet of zero thickness. It is actually very easy to obtain the radiation emitted by a beam of finite thickness for the case of no incoming radiation, i.e., for $B_1=0$. As can be seen from Eqs. (24), (25), and (26) only the amplitude of the emitted field depends on the separation d; the phase ψ_2 is independent of d. Therefore the Poynting vector is given by

$$\langle s_n \rangle = R_0\eta(\epsilon,\beta)\left[\int_0^\infty d\xi (I\alpha)e^{-q\xi}\right]^2 \times\left[\frac{\text{watts}}{\text{cm}^2}\text{ for } I \text{ in }\frac{\text{amp}}{\text{cm}^2}\right] \tag{36}$$

where ξ is measured from the surface of the dielectric.

B. Charge Sheet Sandwiched between Two Dielectrics

In the case considered above, the Čerenkov radiation vanishes for $\beta\to 1$ regardless of how close the electron beam is to the surface. On the other hand, it is well known[4] that a beam going through a dielectric radiates more, the faster the beam moves. In order to understand this difference in behavior, we have to investigate a geometry which contains both these situations as limiting cases. We choose a thin charge sheet between two dielectrics (Fig. 3). The charge density is given by Eq. (1) and the primary fields are Eqs. (13) and

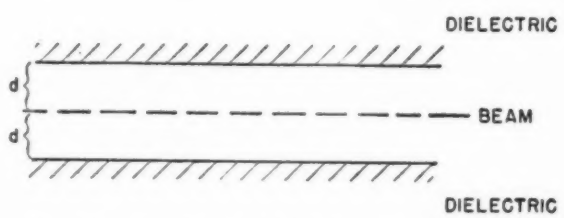


FIG. 3. Flat beam sandwiched between two dielectrics.

(14). The secondary field has to contain terms with both e^{-qx} and e^{qx}. A solution for the vector potential obeying both the wave equation and the Lorentz condition is:

$$\mathbf{A}=(A_1/\vartheta)e^{qx}[\mathbf{z}_0\vartheta\cos(kz-\omega t+\varphi_1]+\mathbf{x}_0\sin(kz-\omega t+\varphi_1)]+(A_2/\vartheta)e^{-qx}[\mathbf{z}_0\vartheta\cos(kz-\omega t+\varphi_2)+\mathbf{x}_0\sin(kz-\omega t+\varphi_2)]. \quad (37)$$

The fields in the dielectric are given by a vector potential of the form in Eq. (19). We shall restrict ourselves to the case with no incoming wave. Therefore in the lower dielectric $B_1=0$ and in the upper dielectric $B_2=0$. The matching conditions in Eq. (23) have to be fulfilled for both $x=d$ and $x=-d$. They yield a set of eight simultaneous equations. Solving for B_1 and B_2 one obtains

$$B_1^2=B_2^2=\frac{4\Phi^2\vartheta^2}{\kappa^2}\frac{(\sinh qd)^2(D^2+1)+1}{(\sinh 2qd)^2(D^2+1)^2+4D^2}. \quad (38)$$

The Poynting vector in the lower dielectric is given by Eq. (29); in the upper dielectric the expression is the same except that $-\mathbf{x}_0$ is replaced by $\mathbf{x}_0$.

Let us now examine the limiting cases. For $\beta\to 1$ both q, Eq. (8), and D, Eq. (28), vanish, and Eq. (38) becomes

$$B_1^2=B_2^2\to\frac{\Phi^2\tau^2}{\epsilon^2\kappa^2} \quad (39)$$

with no dependence on d. Physically this can be understood by building up the secondary and the radiated fields in the following fashion. The primary fields propagate down from $x=0$ to $x=-d$ without phase change. The field strength decreased over this path by e^{-qd}. Here it has to be matched at the vacuum dielectric interface and a secondary field has to be added at the vacuum side. This secondary field now propagates up to the upper interface and loses intensity by e^{-2qd} over this distance. Matching this field at the upper boundary yields a tertiary field, which now is propagated back to the lower interface and so on. In addition there is a field arising from the matching of the primary field at the upper boundary. All these secondary, tertiary etc. fields superimpose to give the secondary field of Eq. (37). The smaller q, the longer the range of the field, and the greater the number of times the field will be "reflected" back and forth between the two surfaces and the greater the amplitude of the secondary field will be with respect to the primary field. As it turns out, the amplitude of the secondary field is given by

$$A_1^2=A_2^2=\frac{\Phi^2\vartheta^2}{\kappa^2}e^{-2qd}\times\frac{(\sinh qd)^2(D^4+1)+D^2\cosh 2qd+1}{(\sinh 2qd)^2(D^2-1)^2+4D^2(\cosh 2qd)^2} \quad (40)$$

which for $\beta\to\lambda$ goes over to

$$A_1^2=A_2^2\to\frac{\Phi^2\tau^2}{4\epsilon^2\kappa^2} \quad (41)$$

and remains finite although the tangential component of the primary field, Eq. (13) vanishes.

In the other limiting case where $e^{-qd}\ll 1$, the expression in Eq. (38) reduces to the no incoming wave case of Eq. (24), as it should because the secondary field originating say at the upper interface has decayed by an additional factor e^{-2qd} before reaching the lower one and its influence is not felt.

C. Cylindrical Geometry

The other of the cases having an easy solution is that of cylindrical symmetry (Fig. 4). Let us introduce a cylindrical coordinate system (r,φ,z), then the channel consists of the space $r<b$ if b is its radius. Let us assume for the charge density

$$\rho=\rho_0\delta(r-a)[1+\alpha\cos(kz-\omega t)]. \quad (42)$$

This corresponds to a charge distribution consisting of a hollow cylinder with a thin wall, and let the charge move with velocity v in the z direction. We now first have to find the primary field, i.e., the field associated with the electron beam which would be the complete solution for $b\to\infty$. The scalar potential of the primary field of the ac part of Eq. (42) is

$$V=V_0\cos(kz-\omega t)K_0(qr) \quad \text{for} \quad r\geq a \quad (43)$$

$$V=V_i\cos(kz-\omega t)I_0(qr) \quad \text{for} \quad r\leq a. \quad (44)$$

The vector potentials are given by Eq. (5). The constants V_0 and V_i have to be determined by the matching conditions at $r=a$. There first has to hold

$$\lim_{\delta\to 0}[V(a+\delta)-V(a-\delta)]=0. \quad (45)$$

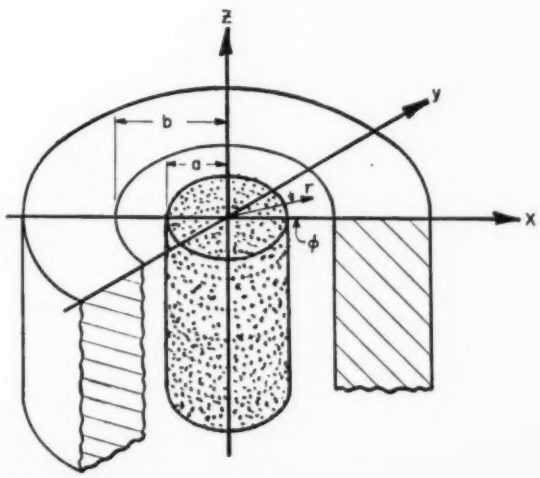


FIG. 4. Hollow beam in cylindrical hole.

The second condition is obtained by performing an integration over r with the wave equation in Eq. (3) as the integrand

$$\int_{a-\delta}^{a+\delta} dr \Box V \equiv \int_{a-\delta}^{a+\delta} dr \left[\frac{\partial}{\partial r} \frac{1}{r} \frac{\partial}{\partial r} (rV) - \left(k^2 \vartheta^2 - \frac{1}{r^2} \right) V \right] = -4\pi \int_{a-\delta}^{a+\delta} dr \rho. \quad (46)$$

The second term on the left-hand side of Eq. (46) disappears when going to the limit $\delta \to 0$ because of Eq. (45) and the second condition is

$$\lim_{\delta \to 0} \left\{ \left[\frac{1}{r} \frac{\partial}{\partial r} (rV) \right]_{r=a+\delta} - \left[\frac{1}{r} \frac{\partial}{\partial r} (rV) \right]_{r=a-\delta} + 4\pi \int_{a-\delta}^{a+\delta} \rho d\sigma \right\} = 0. \quad (47)$$

Using Eqs. (45) and (47) one obtains

$$V_0 = 2a\Phi I_0(qa) \quad (48)$$

$$V_i = 2a\Phi K_0(qa). \quad (49)$$

In order to match the fields at the vacuum dielectric interface $r=b$, one has to add a secondary field of the form of Eq. (44) in the channel and an outgoing wave in the dielectric. The latter is given by the fields

$$\mathbf{E}_D = \mathbf{r}_0 \frac{M}{\epsilon\beta} [\cos(kz-\omega t-\xi) J_1(pr) - \sin(kz-\omega t-\xi) N_1(pr)] + \mathbf{z}_0 \frac{\tau M}{\epsilon\beta} [\sin(kz-\omega t-\xi) J_0(pr) + \cos(kz-\omega t-\xi) N_0(pr)] \quad (50)$$

$$\mathbf{H}_0 = \varphi_0 M [\cos(kz-\omega t-\xi) J_1(pr) - \sin(kz-\omega t-\xi) N_1(pr)]. \quad (51)$$

The matching conditions yield for the constants in Eqs. (50) and (51)

$$M^2 = \frac{f^2}{g^2+h^2} \quad (52)$$

$$\sin\xi = \frac{g}{(g^2+h^2)^{\frac{1}{2}}}; \quad \cos\xi = \frac{h}{(g^2+h^2)^{\frac{1}{2}}} \quad (53)$$

where

$$f = \frac{V_0 D}{b} \quad (54)$$

$$g = DI_0(qb) N_1(pb) + I_1(qb) N_0(pb) \quad (55)$$

$$h = DI_0(qb) J_1(pb) + I_1(qb) J_0(pb). \quad (56)$$

For the time average of the Poynting vector one obtains

$$\langle \mathbf{s} \rangle = \frac{1}{2\pi r} \frac{c}{2\pi} \frac{M^2}{\epsilon\beta k} \left[\mathbf{r}_0 + \frac{1}{\tau} \mathbf{z}_0 \right]. \quad (57)$$

If the charge is distributed over a finite range of r the radiation is given by

$$\langle s_r \rangle = \frac{1}{2\pi r} \frac{c}{2\pi\epsilon\beta k} \left[\int_0^b dr M(r) \right]^2. \quad (58)$$

Setting $b-a=d$ and letting $b \to \infty$ the expression in Eq. (57) goes over into Eq. (29), as it should. In order to obtain the radiation emitted by a single electron going through a channel one has to let a go to zero and ρ_0 go to infinity in such a manner that Eq. (48) assumes the magnitude of the Fourier component of the frequency ω for the single electron.[3] Now the Fourier integrals for $\mathbf{E}$ and $\mathbf{H}$ can be inserted into the expression for the Poynting vector. Integration over the time yields Ginsburg's result.[4]

ACKNOWLEDGMENTS

The author wishes to thank Dr. C. H. Townes, Dr. S. Geschwind, and Mr. H. Lashinsky for many valuable discussions.

APPENDIX

The radiation emitted by a single electron moving parallel to a flat dielectric can be obtained in the following fashion. The charge distribution is

$$\rho = e\delta(z-vt)\delta(x)\delta(y) = \frac{e\delta(x)}{4\pi^2 v} \int_{-\infty}^{+\infty} \int d\omega dl e^{i(ly+kz-\omega t)}. \quad (59)$$

It is unnecessary to perform a Fourier analysis along the x axis since the matching has to be done only at the one value $x=-d$. Similarly the potentials are written as Fourier integrals:

$$V_0 = \frac{e}{2\pi v} \int\int d\omega dl \frac{e^{-g|x|}}{g} e^{i(ly+kz-\omega t)} \quad (60)$$

$$g^2 = k^2\vartheta^2 + l^2. \quad (61)$$

The vector potential follows from Eq. (5). The secondary field has the vector potential

$$\mathbf{A}_1 = \int\int d\omega dl \mathbf{\Lambda}(l,\omega) e^{-gx+i(ly+kz-\omega t)}$$

and the vector potential inside of the dielectric is given by

$$\mathbf{B} = \int\int d\omega dl [\mathbf{x}_0 B_x(\omega,l) + \mathbf{y}_0 B_y(\omega,l) + \mathbf{z}_0 B_z(\omega,l)] e^{i[kz+ly-p(x+d)-\omega t]} \quad (62)$$

$$p^2 = k^2\tau^2 - l^2. \quad (63)$$

The phase factor $\exp(-ipd)$ has been added for the sake of convenience. As can be seen from Eq. (63) the propagation constant in the x direction, p, becomes imaginary for $l^2>k^2\tau^2$, and the corresponding Fourier components become attenuated inside the dielectric. The energy radiated by the electron into the space between z and $z+dz$ is given by

$$\frac{dP}{dz}dz=dz\int_{-\infty}^{+\infty}dt\int_{-\infty}^{+\infty}dy\frac{c}{4\pi}[\mathbf{E}_D\times\mathbf{H}_D]$$

$$=dz\frac{c}{4\pi}\int\int_{-\infty}^{+\infty}dtdy\Big[\int\int_{-\infty}^{+\infty}d\omega dl\mathbf{E}(\omega,l)\times\int\int_{-\infty}^{+\infty}d\omega' dl'\mathbf{H}(\omega',l')\Big].$$

Both $\mathbf{E}_D$ and $\mathbf{H}_D$ depend on y and t only through the exponential in Eq. (62). The integration over these variables therefore merely introduces δ functions:

$$\frac{dP}{dz}=\frac{c}{4\pi}(2\pi)^2\int\int_{-\infty}^{+\infty}\int\int d\omega dld\omega' dl'\delta(\omega+\omega')\times\delta(l+l')[\mathbf{E}(\omega,l)\times\mathbf{H}(\omega',l')]. \quad (64)$$

In carrying out the integration over ω' and l' one has to choose the order of integration and go around the branching points in such a way that P is real. Thus one finally obtains for the energy radiated in the frequency intervals between ω and $\omega+d\omega$ into the space between z and $z+dz$

$$\frac{dP(\omega)}{dz}dzd\omega=dzd\omega 4\pi ck\beta\times\int_0^{k\tau}\frac{dl}{p}[k^2\tau^2|B_y|^2+kl(B_yB_z^*+B_y^*B_z)+(p^2+k^2)|B_z|^2]. \quad (65)$$

In Eq. (65) the component B_x has been eliminated by the Lorentz condition and

$$B_y=-\frac{ele^{-qd}}{\pi vk\beta}\frac{\epsilon g+ip}{\epsilon^2g^2+p^2}$$

$$B_z=\frac{ele^{-qd}}{\pi vk^2\beta}\times\frac{g[k^2\epsilon\tau^2-p^2(\epsilon+1)]+ip[k^2\epsilon\tau^2-p^2(\epsilon+1)+k^2(\epsilon-1)]}{\epsilon^2g^2+p^2} \quad (66)$$

follow from the matching of the fields. The integration over l in Eq. (65) has to be performed only to the upper limit as indicated, since for larger values of l the radiation becomes attenuated and does not contribute to the Poynting vector.

Magnetic Materials for Digital-Computer Components.* I. A Theory of Flux Reversal in Polycrystalline Ferromagnetics†

N. MENYUK AND J. B. GOODENOUGH
Lincoln Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received July 19, 1954)

It is proposed that magnetization reversal in polycrystalline ferro- and ferrimagnetic materials is primarily due to the nucleation and growth of 180° Bloch walls. The origin of domains of reverse magnetization is discussed. The rate of growth of these domains is determined by a study of the elastic and frictional forces which retard the motion of their 180°-Bloch-wall boundaries. This theoretical model successfully explains the output-voltage wave forms of polycrystalline materials. A figure of merit for the magnetization reversal of magnetic cores is defined as the switching coefficient $S_w=(H_m-H_0)\tau$, where τ is the time required to reverse the magnetization, H_m is the applied magnetic field, and H_0 is the threshold field value at which the average domain-wall velocity is zero. S_w is composed of an eddy-current contribution S_w^e and a spin-relaxation contribution S_w^r. The value of S_w is derived in terms of various fundamental parameters of the material. It is shown that in ferrites and ultra-thin metal tapes, $S_w^e \ll S_w^r$. Theoretical relationships expressing the contributions of spin-relaxation, eddy-current, and hysteresis effects to energy losses are derived on the basis of this model. A study of the factors which affect the magnetization-reversal time of materials with square hysteresis loops indicates that a better figure of merit will result if the proper hysteresis shape is obtained by grain alignment rather than by the techniques currently employed in ferrites. A number of experiments are presented in support of this model.

I. INTRODUCTION

A SEARCH for a suitable magnetic material for a small core with (1) a square $B-H$ loop, (2) a short flux-reversal time, τ, (3) a low coercive force, H_c, and (4) a simplicity of manufacture has stimulated a study of the flux-reversal mechanism in polycrystalline materials.

The squareness of a hysteresis loop, which is measured by the squareness ratio $R_s=B_d/B_m$ (see Fig. 1), has been found in many materials to depend upon the degree of alignment of the axes of easy magnetization in the individual grains. These materials show a marked increase in R_s, for instance, after a magnetic anneal, grain orientation, or application of a stress when these treatments tend to align the directions of easy magnetization parallel to the applied field.

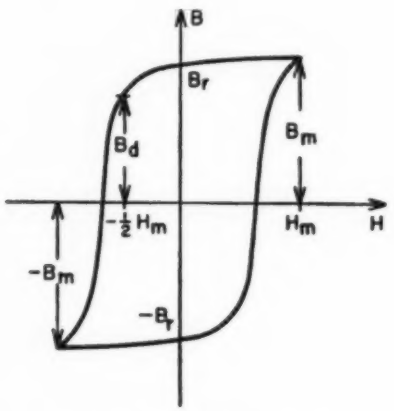


FIG. 1. Definition of $R_s=B_d/B_m$.

The flux-reversal, or switching, time of a magnetic core is defined as the time required to change the induction in a core from its remanent value in one direction to a value corresponding to a driving field strength $H_m>H_c$ in the opposite direction. H_c is the coercivity of the material at operating driving currents. It is not the coercivity of the saturation loop. An irreversible movement of the domain or Bloch walls which separate regions of differently oriented spontaneous magnetization is the principal cause of induction change as the external field strength increases through $H=H_c$. Assume a core is switched by an ideally square driving field H_m and that the 180° domain walls, whose motions account for most of the flux change, are accelerated infinitely rapidly to an average velocity $v=2(\mathbf{H}_m-\mathbf{H}_0)\cdot\mathbf{I}_s/\beta$.[1] H_0 is the threshold field strength at which $v=0$, I_s is the spontaneous magnetization, and β is the viscous-damping parameter for domain-wall motion. If a 180° wall moves with velocity v a distance d in a time τ, then the switching time is

$$\tau=\beta d/[2(\mathbf{H}_m-\mathbf{H}_0)\cdot\mathbf{I}_s]. \quad (1)$$

If the magnetic core is to be used in a magnetic-memory array,[2] $(H_m-H_0)\approx H_c$ at optimum operating conditions. Thus a low value of the coercive force limits the driving pressure which can be applied to the domain walls. Power restrictions, on the other hand, impose an upper limit on H_c. Once an optimum coercive force is attained, therefore, the critical factor in the switching time is $\beta d/I_s$.

The viscous-damping parameter, $\beta=\beta_e+\beta_r$, is composed of two factors. The first arises from the presence

* The research reported in this document was supported jointly by the Department of the Army, the Department of the Navy, and the Department of the Air Force under Air Force Contract No. AF 19(122)-458.
† This paper includes the work reported in Phys. Rev. 91, 434 (1953).

[1] R. Becker, J. phys. radium, 12, 332 (1951).
[2] Jay W. Forrester, J. Appl. Phys. 22, 44 (1951).

of eddy currents. The second is the relaxation contribution which arises from the delayed response of the atomic spins to a force which would change their direction. Whereas β_r is an inherent quality of the magnetic material, β_e is strongly dependent upon the macroscopic dimensions of the sample. Ultra-thin metal tapes of $\frac{1}{8}$-mil thickness have been used to make β_e small. This factor is also small in the ferrites which have resistivities of $\sim 10^6$ ohm-cm.

If the switching time is to be further reduced, the factors which control the magnitude of d and β_r must be known. Since d can be decreased by increasing the number of walls which move irreversibly, Sec. II discusses the creation of domains of reverse magnetization. Since reverse-domain creation alters the induction, there is a discussion of some of the mechanisms which influence the static $B-H$ loop shape. In Sec. III there is a quantitative calculation of τ, β_e, and β_r. The various consequences of the theory are compared with experiment in Sec. V.

II. MAGNETIC-DOMAIN CREATION

The induction in a core is reversed by the motions of many 180° domain walls. The distance any individual domain wall must move will depend upon the proximity of neighboring 180° walls which are moving toward it. If there is a high density of walls present in the material, the distance d of Eq. (1) will be small. New 180° walls are created by the nucleation of domains of reverse magnetization.

However, whenever the magnetization of a region of material is reversed, the total induction through the sample is changed. Therefore, if reverse-domain creation occurs as the external field strength is reduced from H_m to $-H_c$, there will be a "rounding off" of the shoulder of the $B-H$ hysteresis loop. A requirement, then, for a square $B-H$ loop is that no nucleation take place as H decreases from H_m to $-H_c$. The nucleating field strength, H_n, is defined as positive if it is oriented in the direction of the magnetization within the new domain. Therefore, if the distance d is to be reduced compatibly with the requirement of a square $B-H$ loop, the nucleating field strength for any ith domain must satisfy $H_c \leq H_{ni} \leq H_m$.

If $H_{ni} > H_m$, there will be no nucleation of an ith domain to contribute to the switching problem. If $H_{ni} < -H_m$, the only significant field strength for the switching problem is the critical field strength for irreversible growth of the existing ith domain, H_{wi}. If $H_{wi} > H_m$, the ith domain will contribute no mobile 180° walls to influence the distance d of Eq. (1).

Domains of reverse magnetization are created at lattice imperfections. Goodenough[3] has made a theoretical study of the possible lattice imperfections within a magnetic material which can act as nucleating centers for domains of reverse magnetization. It was found that within the range of magnetic field strengths which are used in switch-core and coincident-current memory applications, the grain boundaries are the primary source of mobile 180° domain walls. Domains of reverse magnetization are created in order to reduce the energy which is associated with the grain-boundary surface magnetic-pole density $\omega^* = I_s(\cos\theta_1 - \cos\theta_2)$, where θ_1 and θ_2 are the angles between the magnetization vectors of the neighboring grains and the normal to their common boundary.

The requirement of a positive nucleation field strength H_n, which is necessary for a square hysteresis loop, was shown to be fulfilled provided

$$LI_s^2(\cos\theta_1 - \cos\theta_2)^2 < 60\sigma_w$$

where L is the mean grain diameter and σ_w is the surface energy density of a 180° Bloch wall. This relationship is satisfied in metals if $(\cos\theta_1 - \cos\theta_2)$ is made extremely small. This is commonly done by means of grain orientation, magnetic anneal, or the application of a tensile or compressive stress.

Square loops have also been obtained in polycrystalline ferrites by the application of a compressive stress.[4] However, polycrystalline ferrite materials have been prepared which have square hysteresis loops without any apparent orientation of the axes of easy magnetization from grain to grain.[5,6] Since the magnetization of ferrites is an order of magnitude lower than in metals, this loop squareness is presumably due to the relatively large ratio σ_w/I_s^2.

There are always several factors which contribute to the coercive force of a magnetic material. A study has been made[3] of the contributions which are due to granular inclusions which have closure domains associated with them, lamellar precipitates, grain boundaries, and the domain-wall tension. The granular-inclusion contribution was found to be

$$H_w(\text{incl.}) = C_2(\sigma_w/I_s\langle R\rangle)P^{\frac{1}{2}}$$

where P is the percentage of granular precipitate present, $\langle R\rangle$ is the average inclusion radius, and $C_2 \approx 0.25$ is a dimensionless constant. This equation is valid in metals for $R > 10^{-5}$ cm and in ferrites for $R > 10^{-4}$ cm. If the inclusions are too small for closure-domain formation, their contribution to the coercivity is believed to be considerably reduced.

The lamellar-precipitate contribution to the coercivity is

$$H_w(\omega_l^*) \approx 0.1P'\langle \omega_l^{*2}\rangle / I_s$$

where ω_l^* is the magnetic pole density at the precipitate surface and P' is the percentage of lamellar precipitate present. If an extensive amount of lamellar precipitate is present, the hysteresis loop will not be square.

[3] J. B. Goodenough, Phys. Rev. **95**, 917 (1954).

[4] Williams, Sherwood, Goertz, and Schnettler, Comm. and Elect. **9**, 531 (1953).

[5] E. Albers-Schoenberg, J. Appl. Phys. **25**, 152 (1954).

[6] I. J. Hegyi, J. Appl. Phys. **25**, 176 (1954).

The grain boundaries introduce a coercive-force term

$$H_w(\omega^*) \approx \langle \omega^{*2} \rangle / 2I_s = \tfrac{1}{2} I_s \langle (\cos\theta_1 - \cos\theta_2)^2 \rangle.$$

This contribution is small in ferrites with low I_s and in materials in which the axis of easy magnetization has been aligned from grain to grain.

The model of growing ellipsoidal domains of reverse magnetization gives rise to a coercivity term due to the tension of the Bloch walls. As a reverse domain grows, the area of Bloch wall surrounding the domain increases. The work required to create more Bloch-wall area contributes to the coercive force a term

$$H_w(\sigma_w) = \sigma_w / (2I_s \langle r_0 \cos\theta \rangle)$$

where r_0 is the semiminor axis of an ellipsoid of large eccentricity.

In view of the above, it is concluded that a satisfactory core for coincident-current memory application should be single phase. If any precipitate is present, it must be granular with radius $R < 10^{-5}$ cm in metals and $R < 10^{-4}$ in ferrites.

Since grain boundaries are the nucleating centers for domains of reverse magnetization, it would appear advantageous to reduce the grain size of the material. This would presumably reduce the distance d of Eq. (1). However, a reduction of the mean grain diameter L would also decrease the semiminor axis of the reverse domain r_0, thereby increasing the coercive-force contribution $H_w(\sigma_w)$. In ferrites with low I_s this may be the most significant contribution to the coercivity. Thus, for any given material, there will be an optimum grain size for coincident-current-memory application.

III. THE SWITCHING MECHANISM

A. The Switching Coefficient

The equation of motion of a 180° Bloch wall bounding an expanding, cylindrical domain of radius r and length l can be derived from the Lagrangian and dissipation function.[7] If m is the mass per unit area of the Bloch wall, the kinetic energy is $T = \frac{1}{2} m (2\pi r l) \dot{r}^2$. To expand the domain, work must be done to create a larger area of wall with surface energy density σ_w. Work must also be done against an elastic force per unit area, $\alpha' r$. The potential energy is therefore

$$V = \sigma_w (2\pi r l) + (\pi r^2 l) 2\alpha' r - 2\mathbf{H} \cdot \mathbf{I}_s (\pi r^2 l)$$

so that the Lagrangian becomes

$$L = (2\pi r l) \{ \tfrac{1}{2} m \dot{r}^2 - \sigma_w + (\mathbf{H} \cdot \mathbf{I}_s) r - \tfrac{1}{2} \alpha' r^2 \}.$$

The dissipation function is

$$D = (2\pi r l) \tfrac{1}{2} \beta \dot{r}^2,$$

where the viscous-damping parameter is $\beta = \beta_e + \beta_r$.

[7] P. K. Baltzer, Massachusetts Institute of Technology Digital Computer Laboratory Memorandum M-2275, July, 1953.

If $\alpha = 3\alpha'/2$, the resulting equation of motion is

$$m\left(\frac{1}{2}\frac{\dot{r}^2}{r} + \ddot{r}\right) + \beta\dot{r} + \alpha r + \frac{\sigma_w}{r} = 2(\mathbf{H} \cdot \mathbf{I}_s). \qquad (2)$$

The mass per unit area is $m = \sigma_w / (8\pi\gamma^{*2} A)$ where A is the exchange parameter (see Appendix II). The exchange parameter is defined in Appendix I as $A = 2JS^2/a \approx kT_c/a$ where k is Boltzmann's constant, T_c is the Curie temperature, and a is the lattice parameter.[8] The parameter γ^* is defined in Appendix II. Since $\sigma_w \lesssim 1$ erg/cm^2, $\gamma^{*2} \approx 4 \times 10^{14}$, and $A \sim 10^{-6}$, the mass per unit area of wall is $m \sim 10^{-10}$ g/cm^2. Galt, Andrus, and Hopper[9] have found the inertial term $m\ddot{r} \ll \beta\dot{r}$ in magnetite and nickel ferrite. The output voltage *vs* time curves from the sense winding on metallic and ferrite cores indicate that the domain walls are accelerated to a mean velocity within the rise time of the square input pulse. Rise times of less than 0.2 μsec are used. These experiments are discussed in Sec. IV and V. Such considerations appear to justify the omission of the inertial term $m\ddot{r}$ in Eq. (2). The retarding forces per unit area of the wall can be expressed as energy per unit volume $2\mathbf{H}_0 \cdot \mathbf{I}_s$. H_0 is defined as the threshold field for which $\langle d\langle r \rangle / dt \rangle = 0$. Then Eq. (2) reduces to

$$\tfrac{1}{2} m \langle \dot{r}^2 / r \rangle + \beta \langle \dot{r} \rangle = 2(\mathbf{H} - \mathbf{H}_0) \cdot \mathbf{I}_s. \qquad (2')$$

Measurements[9] of the velocity of a plane wall show $\dot{r}/(H - H_0) \approx 10^3 - 10^4$. Since $r \sim 2 \times 10^{-4}$ cm.

$$\frac{\frac{1}{2} m \langle \dot{r}^2 / r \rangle}{2(\mathbf{H} - \mathbf{H}_0) \cdot \mathbf{I}_s} \approx \frac{(H - H_0)}{10 I_s} \quad \text{to} \quad \frac{10(H - H_0)}{I_s} \ll 1$$

for small applied-field values. The first term of Eq. (2') is negligible in materials characterized by slow wall motion (e.g., magnetite). It can, however, have a small, but measurable, effect in materials characterized by rapid wall motion (e.g., nickel ferrite). To a first approximation this term can be neglected.

In a toroid the wall velocity of a growing, ellipsoidal domain of reverse magnetization is $v = \dot{r}(1 + \lambda^2)^{\frac{1}{2}} / \cos\theta$. Here r is the component of the semiminor axis on the toroid cross section through the center of the domain and λ is the ratio of the semiminor to semimajor axis of the ellipsoidal domain. If $\lambda \ll 1$ and this latter definition of r is used, Eq. (2') becomes

$$\beta \frac{d\langle r \rangle}{dt} = 2 I_s (H_m - H_0) \langle \cos\theta \rangle^2. \qquad (2'')$$

This relation assumes an ideally square input pulse. It will be valid only when the rise time, t_r, of the input pulse is a small fraction of the switching time, τ. Because H_0 represents an average over all the walls which move, it will be a constant value only for driving

[8] C. Kittel, Revs. Modern Phys. **21**, 541 (1949).

[9] Galt, Andrus, and Hopper, Revs. Modern Phys. **25**, 93 (1953).

fields greater than the critical field, $H_w(\max)$, required to move all the domain walls. Equation (2″) should, however, be a good approximation within the limits $H_w(\max)<H_m<H(t_r/\tau<0.3)$.

Kittel[10] has indicated the essential features in the calculation of β_r, the relaxation contribution to the viscous damping parameter $\beta=\beta_r+\beta_e$. Since $\lambda\ll 1$, the velocity of the Bloch walls is essentially that of a cylindrical wall everywhere parallel to the direction of magnetization. Although Kittel's calculation is for a plane wall perpendicular to the z axis moving in the z direction, his formalism holds equally well for a cylindrical wall provided the variable $r/\cos\theta$ replaces the variable z and the integrations are taken from zero to infinity. However, the assumption of a single direction of easy magnetization is implicit in setting $\sigma_w=4[A(K+\lambda_i\sigma)]^{\frac{1}{2}}$ (see Appendix I) where λ_i is the isotropic magnetostriction coefficient and σ is the stress producing internal strains at the Bloch wall. There results (see Appendix III)

$$\beta_r=4\pi\Lambda m=\frac{2\Lambda}{\gamma^{*2}}\left(\frac{K+\lambda_i\sigma}{A}\right)^{\frac{1}{2}}. \tag{3}$$

A study has been made of the velocity of a collapsing cylindrical domain wall in a metal rod.[11] The relaxation effect in this instance is much smaller than the eddy-current effect, and the approximation $\beta=\beta_e$ is valid. The analysis for a cylindrical wall of radius r which is expanding with a radial velocity $\dot{r}$ in a cylindrical specimen of radius r_m, electrical resistivity ρ, and spontaneous magnetization I_s is completely analogous. Since these authors were considering the case for high driving fields only, they equated the power dissipated by eddy currents, A, to the power generated by the alignment of flux with the driving field, B. For low field strengths it is necessary to include the power dissipated in overcoming the potential barrier to wall movement, C. The relation $A+C=B$ together with Eq. (2″) gives (see Appendix IV)

$$\beta_e=\beta_e^0\frac{\langle r\rangle\ln\langle r_m/r\rangle}{\langle\cos\theta\rangle};\quad \beta_e^0=\frac{64\pi^2 I_s^2}{\rho c^2}. \tag{4}$$

In a rectangular specimen, such as a metallic ribbon core, $2r_m$ should be taken as the smaller dimension in the cross section perpendicular to the flux path. Although there may be many small domains which are expanding so that the eddy currents between moving walls tend to cancel one another, there is no cancellation of the eddy currents near the surface of the specimen.

If the finite rise time of the square driving pulse is neglected, (H_m-H_0) is time independent. If d is the distance a wall, moving with the average velocity $d\langle r\rangle/dt$, travels in the time τ, then integration of Eq. (2″) gives

$$(H_m-H_0)\tau=S_w=\beta_r d/(2I_s\langle\cos\theta\rangle)+S_w{}^e, \tag{5}$$

$$S_w{}^e=\frac{\beta_e^0}{8I_s\langle\cos\theta\rangle}\begin{cases}\{2\ln[r_m/(\langle\cos\theta\rangle d)]+1\}d^2 & \text{if } \langle\cos\theta\rangle d\le r_m\\ r_m{}^2/\langle\cos\theta\rangle^2 & \text{if } \langle\cos\theta\rangle d\ge r_m.\end{cases}$$

S_w is defined as the switching coefficient. It is a constant which depends upon the properties of the material in the core. Equation (5) is to be compared to Eq. (1).

Under ideal operating conditions for memory storage, $(H_m-H_0)\approx H_c$. If, therefore, the switching time is to be less than 1 μsec and coercivities of less than 1 oe are to be used, the switching coefficient must be of the order of 10^{-7} oe-sec. The eddy-current contribution, $S_w{}^e$, is made negligibly small by using ferrites or metal tapes of $\frac{1}{8}$-mil thickness.

B. Output Voltage

The output voltage from the sense winding is a measure of the rate of change of flux through the magnetic core. The rate of change of flux in the material is

$$\frac{d\Phi}{dt}=2\frac{d\int\mathbf{B}\cdot d\mathbf{A}_c}{dt}=2\langle\cos\theta\rangle B_s\frac{dA_c}{dt}.$$

The "2" enters because the flux change is the result of a reversal of magnetization in the specimen. The total area of reversed magnetization in the core cross section is $A_c=\sum_i A_i$, where A_i represents the contribution from a single ith domain. According to the model of this paper, the A_i represent cross sections of the domains of reverse magnetization which are either created when $H>H_{ni}$ or are not annihilated by the previous driving field $H=-H_m$. These areas originate in the core cross section at different times and grow at different rates. If n is the total number of A_i which appear and grow on the core cross section when the core is "switched," then $A_c=nf(\langle r\rangle)$ where $\langle r\rangle$ is a function of time. If a distribution function $F(\langle r\rangle)=\partial A_c/\partial\langle r\rangle$ is defined, then the contribution to the voltage output from irreversible wall motion is

$$\left(\frac{d\Phi}{dt}\right)_{irr}=\frac{16\pi I_s^2\langle\cos\theta\rangle^3}{\beta}F(\langle r\rangle)(H_m-H_0). \tag{6}$$

$F(\langle r\rangle)$ measures the rate of change of flux as a function of the constant-velocity-wall positions. Since the average wall velocity is assumed to accelerate to a constant velocity in a time $t\ll\tau$, $F(\langle r\rangle)$ is a maximum when the wall area encompassing the growing domains is a maximum. The wall area surrounding a growing domain increases with r until it collides with similarly growing domains and is annihilated. Because of the random distribution of the growing domains throughout the specimen, at larger values of r, $F(\langle r\rangle)$ decreases to

[10] C. Kittel, Phys. Rev. **80**, 918 (1950).
[11] Williams, Shockley, and Kittel, Phys. Rev. **80**, 1090 (1950).

zero with a Gaussian-like tail. Since $\langle r\rangle$ is a function of time through Eq. (2''), $F(\langle r\rangle)$ is a unimodal function of time. Its maximum shifts to shorter times with increasing H_m. Thus $(d\Phi/dt)_{irr}$ is a unimodal function of time the maximum of which increases in amplitude and shifts to shorter times with increasing applied field strength H_m.

C. Energy Considerations

If a magnetic core is to be used as a high-speed, multiposition switch, it may be operated at frequencies in excess of 5 kilocycles. If energy losses are excessive, sufficient heat will be generated to raise the cores to a temperature at which the satisfactory magnetic characteristics are destroyed. Under the assumptions inherent in Eq. (2'')

$$\langle\beta_e v\rangle_t+\langle\beta_r v\rangle_t+\langle 2\mathbf{H}_0\cdot\mathbf{I}_s\rangle_t=2\mathbf{H}_m\cdot\mathbf{I}_s.$$

The energy dissipations per unit volume per half cycle due to eddy currents, relaxation effect, and hysteresis are, respectively,

$$\langle\beta_e v\rangle_t=\frac{1}{\tau}\int_0^\tau\beta_e v\,dt=(2I_s\langle\cos\theta\rangle S_w^e)/\tau,$$

$$\langle\beta_r v\rangle_t=\frac{1}{\tau}\int_0^\tau\beta_r v\,dt=\beta_r d/\tau,$$

$$2\langle\mathbf{H}_0\cdot\mathbf{I}_s\rangle_t\approx 2H_0I_s\langle\cos\theta\rangle$$

provided the hysteresis loop is square. $2\mathbf{H}_m\cdot\mathbf{I}_s$ is the energy per unit volume supplied by an ideally square-wave external field of amplitude H_m. The total energy loss per unit volume per cycle is, therefore,

$$\frac{1}{\tau}\{4I_s\langle\cos\theta\rangle S_w^e+2\beta_r d\}+4H_0I_s\langle\cos\theta\rangle. \tag{7}$$

D. Comparison of Preparation Techniques

In an ultra-thin metal tape or a ferrite in which the electrical resistance is very large $S_w^e\approx 0$, and Eq. (5) reduces to

$$(H_m-H_0)\tau=S_w=\beta_r d/2I_s\langle\cos\theta\rangle. \tag{5'}$$

Since the switching coefficient S_w is a figure of merit for the core which should be as small as possible, the distance d should be as small as possible compatible with a square $B-H$ loop. If the mechanism for reverse-domain creation outlined in Sec. II is correct, any increase in H_n will increase d. The optimum operating point, therefore, is for a square $B-H$ loop with H_n and H_c approximately equal. Of the three methods which might be available for the control of H_n, it is of interest to determine which yields the fastest switching time τ.

According to Sec. II the threshold field H_0 may be written

$$H_0=H_0'+\xi I_s\langle(\cos\theta_1-\cos\theta_2)^2\rangle+\eta\sigma_w/I_s$$

where ξ and η are functions which are independent of I_s, σ_w, and $(\cos\theta_1-\cos\theta_2)$. Actually, $\eta\propto\langle r_0\rangle^{-1}$. It will be assumed, however, that a variation in H_n will have a relatively small effect on r_0. This assumes that a decrease in ω^* will decrease the number of domains which are created without altering their mean size. Similarly S_w is proportional to $\sigma_w d/I_s$ where the constant of proportionality ζ is a function like ξ and η. Therefore the switching time may be written

$$\tau=Qd;$$

$$Q=\frac{\zeta\sigma_w/I_s}{(H_m-H_0')-\xi I_s\langle(\cos\theta_1-\cos\theta_2)^2\rangle-\eta\sigma_w/I_s}. \tag{8}$$

The three methods which are available for increasing H_n sufficiently to create a square $B-H$ loop are (1) reduction of $(\cos\theta_1-\cos\theta_2)$ by grain orientation or magnetic anneal, (2) reduction of $(\cos\theta_1-\cos\theta_2)$ by application of a tensile or compressive stress, and (3) increase of σ_w/I_s by a change in chemical composition. It is assumed that H_0', ξ, η, and ζ are not affected by any of the three treatments and that d depends upon the value of H_n. It is shown below that Method (1) yields the shortest value of τ.

If Method (1) is used, $\xi I_s\langle(\cos\theta_1-\cos\theta_2)^2\rangle$ decreases while σ_w/I_s is unchanged. For a given driving field, this reduction in H_0 decreases the coefficient Q of Eq. (8). In metals the reduction in H_0 can be as large as 1 oe.[3] If Method (2) is used, however, there is an increase in $\sigma_w\propto(K+\lambda_s\sigma)^{\frac{1}{2}}$ (see Appendix I) which offsets the decrease in $\langle(\cos\theta_1-\cos\theta_2)^2\rangle$. In the ferrites the decrease of the grain-orientation contribution to the coercive force is less than 0.1 oe and is overshadowed by the change in $\lambda_s\sigma$. Therefore H_0 increases, and the denominator of Q decreases while its numerator is increased. In this case the coefficient Q is increased for any given driving field H_m. If Method (3) is used, σ_w/I_s will be the same as in Method (2) if the same value of H_n is attained. The reduction in $I_s\langle(\cos\theta_1-\cos\theta_2)^2\rangle$ will be accomplished, however, by a reduction in I_s rather than a reduction in $\langle(\cos\theta_1-\cos\theta_2)^2\rangle$.

It is concluded, therefore, that a material, if prepared by Method (1), will have a shorter switching time for a given driving field strength H_m than it would if it were given a square $B-H$ loop by Method (2) or (3). The thin metal tapes are prepared by Method (1). The ferrites are made square by Method (2) or (3).

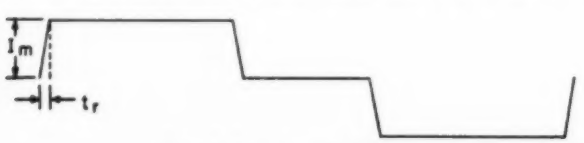


FIG. 2. Input current pulse.

IV. EXPERIMENTAL EQUIPMENT

In a determination of the switching coefficient, the current-input signals to the cores are of the form shown in Fig. 2. The pulse length is greater than the switching

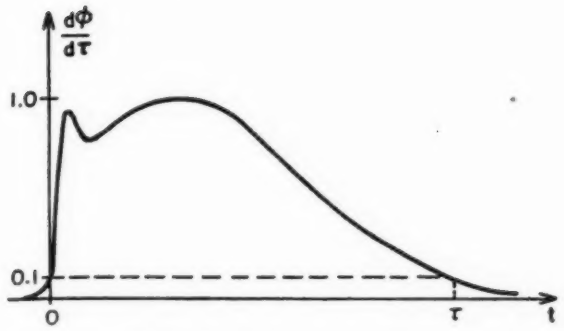


FIG. 3. Definition of switching time.

time of the magnetic core, so for $t_r \ll \tau$, the input signal may be considered a semi-infinite step function. In this experiment, $t_r \approx 0.2$ μsec.

In this laboratory, time is measured from the instant the output voltage from the sense winding of the magnetic core has risen to 10 percent of its maximum amplitude. If the output voltage has a double maximum, τ is the time required for the signal to decay to 10 percent of the voltage at the second maximum (see Fig. 3). This convention gives consistent results provided the total volume of material which reverses its flux is independent of the driving field over the region considered. Since this requirement is fulfilled in materials with a square hysteresis loop, the established convention for measuring switching time was used.

A block diagram of the circuit used in this measurement is shown in Fig. 4. The current, voltage, and time calibrators were designed by B. Gurley of this laboratory. The current calibrator permitted a direct determination of the input current, from which the applied magnetic field H_m was calculated. The voltage calibrator was so designed that a horizontal line was superposed on a Type 514D Tektronix Oscilloscope with the output voltage as taken from the sense winding. The horizontal line was then adjusted until it was at the level of the maximum amplitude of the output voltage. The value of this voltage could then be read directly from a voltmeter.

The horizontal line could then be switched to one-tenth its original amplitude. The distance between the points at which this line crossed the voltage output curve then represents the switching time of the core. This determination was performed by the time calibrator, which consists of a variable pulse delay unit which is calibrated directly in microseconds.

V. COMPARISON WITH EXPERIMENT

A. Shape of the Voltage Output *vs* Time Curve

If a maximum driving field strength of ~2 oe is used to switch a typical magnetic core, little flux change can be attributed to the rotation of elementary magnetic moments against the large internal anisotropy forces. The principal source of flux change is domain creation and Bloch-wall motion. Each of these processes may be either reversible or irreversible. If a wall is moved reversibly by an external driving field, it will return to its original position of minimum potential energy when the field is removed. If a wall is moved irreversibly, it will settle into some new position of minimum potential energy when the driving field is removed.

Materials with a square $B-H$ loop characteristically exhibit, if driven by fields only slightly greater than the coercivity, a double maximum in the output voltage *vs* time curve (see Fig. 3). The rise time to the first maximum has been found to be limited by the rise time of the input pulse. These rise times have been as short as 0.05 μsec. The second maximum, on the other hand, occurs only for driving field strengths $H_m > H_c$. It increases in amplitude and moves to shorter times as the driving field is increased. At large driving field strengths the resolution between the first and second maximum is lost. Ferrites with $H_c > 1$ oe or with $B-H$ loops which are not square show little or no trace of the first maximum.

The output voltage of the sense winding is a measure of the rate of change of flux in the magnetic core. Since the principal maximum of the output voltage *vs* time curve occurs only when $H_m > H_c$; it is proposed that this maximum is due to the irreversible wall motion of many growing domains. According to the model of Sec. III, this maximum should then (and it does) obey the qualitative predictions for the unimodal curve discussed in connection with Eq. (6). The characteristic first maximum, on the other hand, is attributed to domain creation and reversible wall motion. Although these latter effects are present in the ferrites with high coercive force and nonsquare $B-H$ loops, the

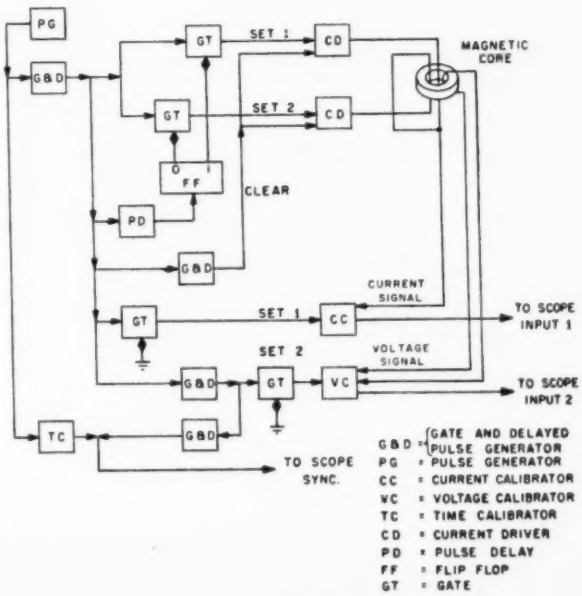


FIG. 4. Block diagram of S_w measurement equipment.

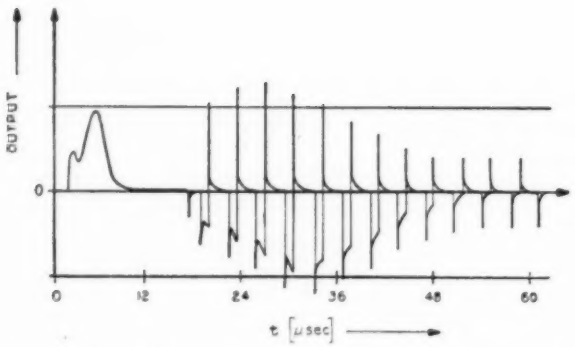


FIG. 5. Output voltage from a rolled, ¼-mil 4–79 Mo-Permalloy ribbon core driven by one long pulse followed by a series of opposite 0.7-microsecond pulses.

resolution of the two maxima is apparently destroyed. The switching times in the materials with a single maximum voltage output are relatively short. In order to substantiate this proposal, three experiments will be discussed.

(1) Rising[12] drove a 4–79 Mo-Permalloy core rolled to ⅛-mil thickness with a series of square pulses, equal in amplitude and sufficient to switch the core. The first pulse was longer than the switching time τ. Succeeding pulses were in the opposite direction and of only 0.7-μsec duration. The resulting oscilloscope trace is reproduced in Fig. 5. The output voltage for the first, long pulse shows the usual double maximum. The second maximum can be nearly reproduced if, once the spikes due to reversible motions are subtracted, the small "bits" from the successive 0.7-μsec pulses are pieced together. The second maximum is therefore dependent upon the state of magnetization of the specimen, or upon $F(\langle r \rangle)$, as predicted by theory. This would not be the case if the second maximum were the result of an increase in the velocity of domain walls with fixed area. The "bits" from the successive 0.7-μsec pulses are not perfectly additive because the walls, at cutoff, settle into the nearest position of minimum potential energy.

The sharp peaks which occur at rise time and cutoff are apparently due to reversible flux changes. They are seen to correspond directly to the first maximum in the uninterrupted voltage output *vs* time curve. If acceleration of the irreversibly moving walls to some average velocity occurs within the rise time of the pulse, as predicted by theory, the rise to the beginning of each bit can be attributed to acceleration of the irreversibly moving walls. The additional sharp peaks are then due to reversible effects.

At cutoff the reversibly moved walls return to their initial position. Since the reversible contributions are nearly constant for each rise time, this contribution can be assumed constant for each cutoff. There is also a contribution to the sharp peak at cutoff from the irreversibly moving walls. If a wall, at cutoff, is moving away from a potential maximum towards a potential minimum, it will continue to move forward to its position of equilibrium within the fall time of the driving pulse. If the wall is moving away from a potential minimum towards a potential maximum, however, it will return to the minimum only after it has first stopped and reversed its direction of motion. Besides the deceleration of the irreversibly moving walls, therefore, there will also be some backward motion. This backward motion contributes to the reverse peak at cutoff. The magnitude of the contribution these irreversibly moving walls make to the cutoff peak will depend directly upon the area of irreversibly moving wall which moves backward to a potential minimum. Accordingly the strongest cutoff peak occurs at the cutoff before the maximum bit. Collisions between the irreversibly moving walls occur during the maximum bit and reduce the wall area which moves backward at cutoff.

(2) If the first maximum is due partially to the nucleation of domains of reverse magnetization, it should be reduced if a complicated domain pattern could be pre-established in an otherwise identical material. Samples of ¼-mil 4–79 Mo–Permalloy ribbon were, respectively, cut in the direction of grain orientation and at 45 degrees to this direction. The former sample, after being pulsed, contains a simple domain pattern. The latter, however, contains a complex pattern as a result of the surface poles which exist. When these cores are pulsed again in the opposite direction, nuclei must be created in the first core whereas they are already present in the second. Figure 6 shows the corresponding output-voltage curves. The first maximum has been greatly reduced in the specimen cut at 45 degrees to the easy magnetization direction despite the initial presence of more domain walls. Thus domain creation, as well as the reversible wall motions, appears to contribute to the first maximum. It is also interesting to note that $S_w = 4.8 \times 10^{-7}$ oe-sec for the 45-degree core as against 6.3×10^{-7} oe-sec for the other. The distance d has apparently been reduced by a factor of 1.3 in the 45-degree core as a result of the more complex domain pattern.

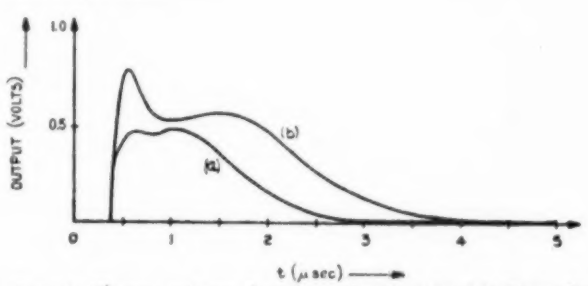


FIG. 6. Output voltage from two ¼-mil 4–79 Mo-Permalloy ribbon cores driven by square-wave pulse of amplitude 0.3 oersted. (a) ribbon cut 45° to rolling direction. (b) ribbon cut parallel to rolling direction.

[12] H. K. Rising, Digital Computer Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts (unpublished research).

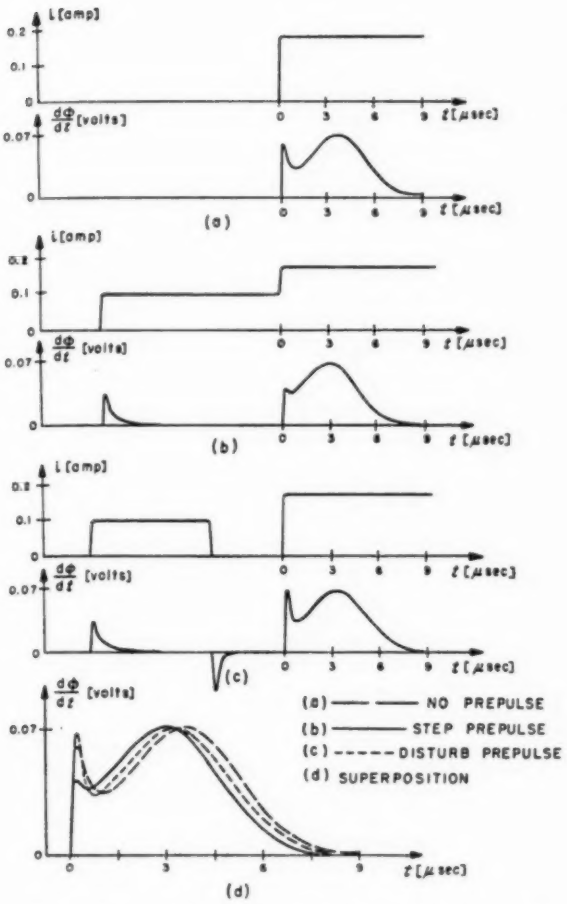


FIG. 7. Voltage output from a ½-mil 4–79 Mo-Permalloy ribbon core after various prepulse conditions.

(3) In order to separate that portion of the first maximum which is due to irreversible domain creation from that which is due to reversible effects, a metal and a ferrite core were subjected to a step prepulse and a disturb prepulse.§ The oscilloscope traces of the driving-current forms and the resulting output voltages are shown in Figs. 7 and 8. After the step prepulse, the first maximum, in each case, is markedly reduced. (The reversible effects should be at best halved by the step prepulse.) Also, there is a shift of the principal maximum to shorter times. Such a shift is predicted by the Sec. III model of many growing domains of reverse magnetization. The average initial radius of a reverse domain at a cross section is larger after the prepulse. Therefore the walls move a shorter distance before collisions occur between the various growing domains. The shift of the maximum voltage output along the time axis is smaller than the rise time of the input pulse. Therefore, this measurement cannot be used to give a quantitative estimate of the variation of initial radii since the velocity of the wall over the distance Δr is limited by the rise time of the input current.

§ The authors would like to thank J. Childress for making these measurements.

After a disturb prepulse is complete, the reversible wall motions are returned to their initial positions. What changes occur in the output-voltage wave form, therefore, must be due to irreversible effects which occur during the prepulse. In the metal core there is a shift of the principal maximum to shorter times and a narrowing and increase of the first maximum. These effects are understandable in terms of a small, irreversible growth, during the disturb pulse, of some of the existing reverse domains. In the ferrite cores, however, there is a shift to longer times of the principal maximum. Within the framework of the model of Sec. III, this is only explainable in terms of irreversible domain creation. If some domains are irreversibly created during the disturb prepulse, their pre-existence will increase the field strength required to create additional reverse domains at grain boundaries. Fewer domains of reverse magnetization will be created after a disturb prepulse than if there were no prepulse. Fewer domains will grow to switch the core, and the domain walls must move farther before collision takes place. That irreversible domain creation occurs in the ferrite and not in the metal is reasonable. The ferrite is prepared

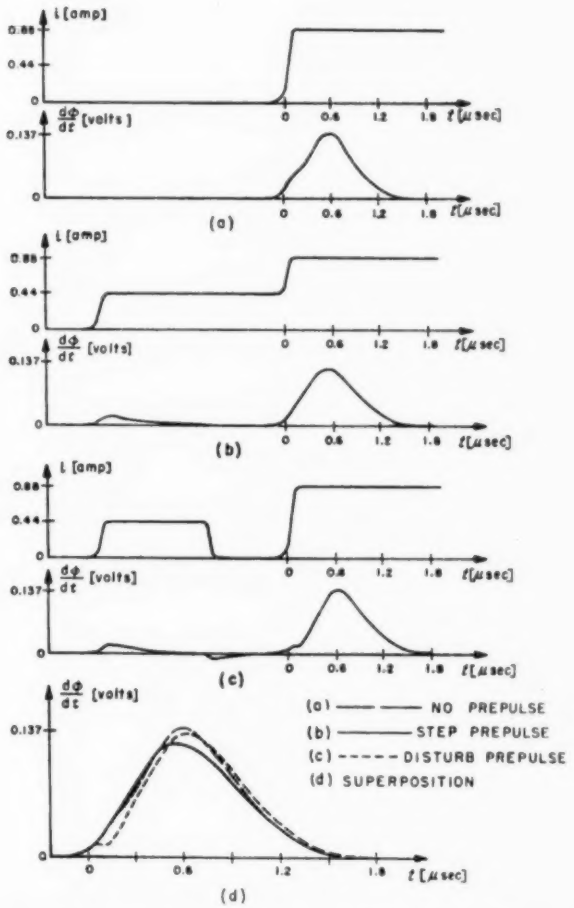


FIG. 8. Voltage output from a General Ceramics MF-1326B magnesium-manganese ferrite core after various prepulse conditions.

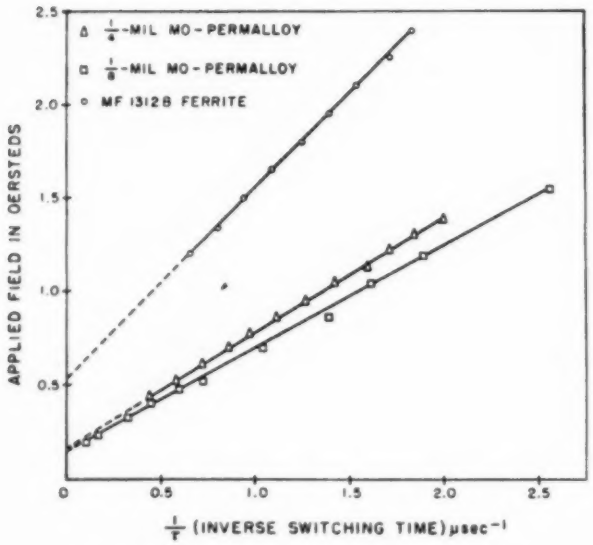


FIG. 9. Measurements of the switching coefficient.

by a sintering of the metal oxides and is full of voids. These inclusions hinder wall motion (see Sec. II) and trap some of the new domain walls which are created. This mechanism is not present in the metal tape. The difference in flux change at the rise time and cutoff of the disturb pulse indicates considerable irreversible change during the prepulse in support of this interpretation.

B. The Switching Coefficient and Energy Losses

The switching coefficient, S_w, has been measured for both ferrites and ultra-thin metal tapes. It is given directly by the slope of the H_m *vs* $1/\tau$ plot, and a straight-line relationship was found to hold for values of H_m within the limits specified in the assumptions of Eq. (2″). In Fig. 9 experimental curves for two 4–79 Mo-Permalloy ribbons rolled respectively to ¼-mil and ⅛-mil thickness are shown. A similar curve for a General Ceramics Co. MF-1312-B ferrite body is also shown. From these curves the measured and theoretical values for S_w can be compared. The energy loss per cycle can be calculated from Eq. (7) and compared with the experimental value.

(1) For both the ¼-mil and ⅛-mil metal ribbon, $H_0=0.14$ oe. The measured values of S_w, however, are 6.3×10^{-7} oe-sec and 5.5×10^{-7} oe-sec, respectively. Since the two samples differ only in their thickness, the difference in S_w should be due to a difference in the eddy-current effects, or in $S_w{}^e$. Since $r_m\leq\langle\cos\theta\rangle d$ and differs by a factor of two between the samples, Eq. (5) gives for the eddy-current contribution to the ⅛-mil-ribbon switching parameter

$$S_w{}^e(\text{exp})=\tfrac{1}{3}(6.3-5.5)\times10^{-7}=0.27\times10^{-7}\text{ oe-sec.}$$

Thus the eddy-current damping is found to be only 5 percent of the relaxation damping in the ⅛-mil metal tapes.

For 4–79 Mo-Permalloy I_s equals 700 (e.m.u.). Since the tape has been grain oriented, $\langle\cos\theta\rangle\approx1$. A tape thickness of ⅛ mil gives $r_m=1.6\times10^{-4}$ cm. The resistivity is $\rho\approx6\times10^{-17}$ statohm-cm. Therefore

$$S_w{}^e(\text{cal})=0.3\times10^{-7}\text{ oe-sec.}$$

To calculate the relaxation-effect contribution to the switching parameter, it is noted that the exchange parameter is $A\sim10^{-6}$ erg/cm and therefore $(K/A)^{\frac{1}{2}}\approx10^5$ cm^{-1}. The magneto-mechanical ratio is $\gamma=1.76\times10^7$ (gauss-sec)$^{-1}$. Although the relaxation frequency Λ is not known for 4–79 Mo-Permalloy, it has been measured as $\Lambda\cong2\times10^8$ sec^{-1} in Supermalloy.[13] In order that $S_w{}^r(\text{cal})=S_w{}^r(\text{exp})=5.2\times10^{-7}$ oe-sec, it is necessary to choose $d=4.8\times10^{-3}$ cm. Since d is the distance a domain wall travels in the time τ and it is believed that at some grain boundaries no domains of reverse magnetization are nucleated, the distance d would reasonably be of the order of a mean grain diameter. Thus the value of d which gives agreement between theory and experiment is as reasonable a one as could have been chosen.

(2) The resistivity of the MF-1312-B ferrite body is so large that the eddy-current contribution to the switching parameter of this material may be neglected. The measured value, therefore, is

$$S_w{}^r(\text{exp})=S_w(\text{exp})=1.02\times10^{-6}\text{ oe-sec.}$$

This body has an $I_s\approx155$ (e.m.u.). If $d=5\times10^{-3}$ cm is again taken, then since the other parameters are approximately the same as for 4–79 Mo-Permalloy, $S_w{}^r(\text{cal.})\sim10^{-6}$ oe-sec.

The theory leading to Eq. (5) has predicted both the straight-line relationships of Fig. 8 and the correct orders of magnitude for the two contributions to the slope of the line, $S_w{}^e$ and $S_w{}^r$. Also, the theory appears to be valid for the ferrimagnetic ferrite as well as for the ferromagnetic Permalloy. Finally, it should be noted that $S_w{}^r$ does not differ markedly in magnitude in the ferrite from its value in the metal. The decreased switching time of the ferrites in memory core applications is the result of higher coercivities.

(3) If the measured values $\tau=1.1$ μsec, $H_0=0.52$ oe, and volume$=0.145$ cm^3 for the MF-1312-B body are used in Eq. (7), a relaxation loss per cycle of 62 ergs and a hysteresis loss per cycle of 44 ergs are calculated. The experimental energy losses per cycle were found to be 87 ergs and 75 ergs respectively. The discrepancy of a factor of two in the hysteresis loss as calculated from that as measured from a static $B-H$ loop is not surprising since this calculation is based on an ideally square $B-H$ loop.

VI. SUMMARY

It has been shown that the output-voltage wave form from the sense winding of a memory core can be

[13] C. Kittel, J. phys. radium **12**, 291 (1951).

explained by a model which attributes the flux change in the core to the creation and subsequent growth of many domains of reverse magnetization. These domains are believed to be predominantly created at grain boundaries. The model assumes that the reverse domains are ellipsoids of large eccentricity. As the domain walls move through the core, their area increases until they start to collide with one another. The changing wall area gives a voltage output of varying amplitude even if the walls move at a constant velocity.

A fundamental parameter for the switching problem has been defined. It is the switching parameter $S_w = S_w^r + S_w^e$. For thin metal tapes and ferrites $S_w^e \ll S_w^r \propto \Lambda\sigma_w d/(I_s T_c)$, where Λ is an intrinsic relaxation frequency and T_c is the Curie temperature. The distance d is approximately equal to a grain diameter. A small grain size, however, also increases the coercivity of the material. It is necessary to determine the optimum grain size for any given material. This is an important feature in determining the firing cycle for ferrites. The factors which influence Λ are not yet understood. A high Curie temperature is desirable both to reduce S_w^r and to insure constant core characteristics when high-frequency operation is required. The factor σ_w/I_s should be kept as low as possible compatible with the loop-squareness requirement $60\sigma_w/I_s^2 > L(\cos\theta_1 - \cos\theta_2)^2$. It is better, therefore, to produce loop squareness by grain orientation or magnetic anneal than by application of an external stress or by a variation of the chemical composition.

The model of this paper has been shown to apply equally well to a ferromagnetic metal-tape core and a ferrimagnetic ferrite core. The switching coefficient of the metal-tape core was found to be smaller by a factor of two than that for the ferrite. The faster switching times of the ferrite memory core is therefore found to result solely from their higher coercivities. In the memory-core application the driving currents are restricted to $H_m \approx 2H_c$, where H_c is the coercivity of the operating loop, not of the saturation loop.

The authors would like to express their thanks to the Magnetic Materials Group Staff for their lively interest in this work and their helpful cooperation in the assembly and operation of the measuring equipment.

APPENDIX I

The calculation of σ_w for a cylindrical wall is similar to that given by Kittel[8] for a plane wall. If only interactions between nearest neighbors are considered, and if these interactions are all equal, the change in exchange energy due to the presence of a wall in which neighboring spins make a small angle $\phi_{ij} < 1$ is

$$\Delta W_{ex} \approx JS^2 \sum_{i<j} \phi_{ij}^2$$

where J is the exchange integral and S is the total spin quantum number. J is so defined that for two spins $\frac{1}{2}$ the energy for parallel orientation is lower than that for antiparallel orientation by an amount $2J$. If the crystal has only one axis of easy magnetization, the small angle can be expanded as

$$\phi_{ij} = r_{ij}\left(\frac{\partial\phi_{ij}}{\partial r}\right) + \tfrac{1}{2}r_{ij}^2\left(\frac{\partial^2\phi_{ij}}{\partial r^2}\right)$$

where r_{ij} is the component of the near-neighbor distance along the radius from the axis of the cylindrical domain to the ith atom. If the axis of the cylindrical domain is along a cubic axis, and if θ is the angle between the radius to the ith atom and a radius parallel to a cubic axis in a b.c.c. lattice, the ith atom has four near neighbors with $|r_{ij}| = [a/(2)^{\frac{1}{2}}]\cos(\pi/4 - \theta)$ and four with $|r_{ij}| = [a/(2)^{\frac{1}{2}}]\cos(\pi/4 + \theta)$ where a is the length of a cube edge. Higher powers of the expansion can be neglected. Then

$$\sum_{i<j} \phi_{ij}^2 = \sum_{i<j} r_{ij}^2\left(\frac{\partial\phi}{\partial r}\right)^2 = 2a^2\left[\cos^2\left(\frac{\pi}{4} - \theta\right) + \cos^2\left(\frac{\pi}{4} + \theta\right)\right]\left(\frac{\partial\phi}{\partial r}\right)^2.$$

There are two atoms per unit cell, and interactions must not be counted twice. The exchange energy density therefore is

$$f_{ex} = \frac{\Delta w_{ex}}{a^3} = A\left(\frac{\partial\phi}{\partial r}\right)^2; \quad A = 2JS^2/a.$$

Since a single axis of easy magnetization has been assumed, the anisotropy energy density is $f_k \approx K\sin^2\phi$. If there are internal strains present, there is also a magnetoelastic term to be considered. If λ_i is the isotropic magnetostriction constant and σ is the stress which causes the internal strains, then the magnetoelastic energy density is $f_{me} = (\frac{3}{2})\lambda_i\sigma\sin^2\theta$ where θ is the angle between the magnetization and the stress.[8] If factors of the order of unity are neglected, $f_K + f_{me} = (K + \lambda_i\sigma)\sin^2\phi = g(\phi)$. The surface energy density of a cylindrical wall of radius r is, therefore,

$$\sigma_w = \frac{1}{r}\int_0^\infty (f_K + f_{me} + f_{ex})r'dr' \approx \int_0^\infty \left[g(\phi) + A\left(\frac{\partial\phi}{\partial r'}\right)^2\right]dr'.$$

The assumption has been made that r is much greater than the wall thickness over which $(f_K + f_{me} + f_{ex}) \neq 0$. Minimization of the integral gives $f_K + f_{me} = f_{ex}$ and therefore

$$\sigma_w \approx 2[A(K + \lambda_i\sigma)]^{\frac{1}{2}}\int_0^\pi |\sin\phi|\,d\phi = 4[A(K + \lambda_i\sigma)]^{\frac{1}{2}}.$$

APPENDIX II

Becker[14] has shown that Bloch-wall motion may be thought to be due to a perpendicular field H_e. The precessional velocity of the spins in the wall as it moves through the lattice is, then,

$$\frac{d\phi(r)}{dt}=-\dot{r}\frac{\partial\phi}{\partial r}=\gamma^* H_e,$$

$$\gamma^{*2}=\gamma^2+\Lambda^2/I_s^2,$$

where ϕ is an angle measuring the spin direction, $\gamma=ge/2mc$ is the magneto-mechanical ratio, and Λ is the intrinsic relaxation frequency defined by the phenomenological equation of motion[15]

$$\frac{d\mathbf{I}}{dt}=\gamma(\mathbf{I}\times\mathbf{H})+\Lambda(\mathbf{I}\times\mathbf{H}\times\mathbf{I})/I^2.$$

If $\delta\sigma_w$ is the difference in wall energy per unit area between the wall in motion and at rest, the kinetic energy of the wall is

$$\tfrac{1}{2}m(2\pi rl)\dot{r}^2=T=(2\pi rl)\delta\sigma_w$$

$$=\left(\frac{2\pi rl}{8\pi}\right)\cdot\frac{1}{r}\int_0^\infty H_e^2 r'dr'$$

and, with the help of Appendix I,

$$\frac{1}{2}m\dot{r}^2=\frac{\dot{r}^2}{8\pi\gamma^{*2}}\frac{1}{r}\int_0^\infty\left(\frac{\partial\phi}{\partial r'}\right)^2 r'dr'=\frac{\dot{r}^2}{8\pi\gamma^{2*}}\left(\frac{\sigma_w}{2A}\right).$$

Therefore the mass per unit area of cylindrical Bloch wall is $m=\sigma_w/8\pi\gamma^{*2}A$.

APPENDIX III

If the applied field, H_m, is parallel to a moving Bloch wall, the total field is $\mathbf{H}=\mathbf{H}_m+\mathbf{H}_e$, where $\mathbf{H}_e$ is the Becker field perpendicular to the Bloch wall and $\mathbf{H}_m$. The power dissipated per unit volume by the motion of the wall is (see Appendix II)

$$\mathbf{H}\cdot\frac{d\mathbf{I}}{dt}=\Lambda H_e^2.$$

The total power required to keep a cylindrical wall moving with a speed $\dot{r}$ is, therefore,

$$(2\pi l)\Lambda\int_0^\infty H_e^2 r'dr'=(2\pi rl)4\pi\Lambda m\dot{r}^2.$$

If the power dissipated is equated to $2(\mathbf{H}-\mathbf{H}_0)\cdot\mathbf{I}_s\times d(\pi r^2 l)/dt$, the power supplied, there results

$$\frac{1}{\dot{r}}=\frac{4\pi\Lambda m}{2(\mathbf{H}-\mathbf{H}_0)\cdot I_s}=\frac{\beta_r}{2(\mathbf{H}-\mathbf{H}_0)\cdot I_s}.$$

APPENDIX IV

If a cylindrical wall of radius r expands in a cylindrical specimen of radius r_m, the voltage induced at r', when $r\leq r'\leq r_m$, is

$$2\pi r'E_\theta=-\frac{1}{c}\frac{d\Phi}{dt}=-\frac{16\pi^2}{c}I_s r\dot{r}.$$

The power dissipated by eddy currents is, therefore,

$$A=2\pi l\int_r^{r_m}\mathbf{j}\cdot\mathbf{E}r'dr'=\frac{2\pi l}{\rho}\int_r^{r_m}E_\theta^2 r'dr'$$

$$=(2\pi rl)\beta_e^0 r\ln\left(\frac{r_m}{r}\right)\dot{r}^2.$$

The power released by magnetization change is

$$B=2\mathbf{H}\cdot\mathbf{I}_s d(\pi r^2 l)/dt,$$

and the power dissipated in moving against the surface tension in the wall and the elastic effects is

$$C=2\mathbf{H}_0\cdot\mathbf{I}_s d(\pi r^2 l)/dt.$$

The balance $A+C=B$ gives

$$\frac{1}{\dot{r}}=\frac{\beta_e^0 r\ln(r_m/r)}{2(\mathbf{H}-\mathbf{H}_0)\cdot\mathbf{I}_s}=\frac{\beta_e}{2(\mathbf{H}-\mathbf{H}_0)\cdot\mathbf{I}_s}.$$

[14] R. Becker, J. phys. radium **12**, 332 (1951).
[15] L. Landau and E. Lifshitz, Physik. Z. Sowjetunion **8**, 153 (1935).

JOURNAL OF APPLIED PHYSICS VOLUME 26, NUMBER 1 JANUARY, 1955

Analysis of Turbulent Density Fluctuations by the Shadow Method

MAHINDER S. UBEROI* AND LESLIE S. G. KOVASZNAY
Department of Aeronautics, The Johns Hopkins University, Baltimore, Maryland
(Received January 18, 1954)

Shadow pictures of turbulent flow regions give a particular kind of information about the three-dimensional density field. If homogeneity and isotropy of the fluctuation field can be assumed, the statistical properties of the picture uniquely determine the correlation function or power spectrum of the three dimensional turbulent density field. Shadow pictures obtained in the wake of a supersonic projectile were analyzed in this manner and encouraging results were obtained. In order to check the validity of the method, the density fluctuations in a heated jet were measured both by the shadow method and by the more conventional hot-wire technique.

I. INTRODUCTION

MEASUREMENT of turbulence usually involves the determination of statistical averages of the fluctuating velocity, pressure, density, and temperature fields. The statistical quantities to be measured can be probability densities or joint probability densities; however, in most cases one is satisfied to obtain information only about the second moments of joint probability densities. These are called the correlation functions. Often the complementary description by power spectra is more convenient. In all cases the information that the experimenter is seeking is of a statistical nature.

The hot wire is almost the unique instrument for measuring turbulent velocity fluctuations in the low-speed regime. With the increase of Mach number, however, the density, pressure, and temperature fluctuations become more pronounced, and the fluctuation field can be decomposed into three "modes" identified in the limit as "velocity mode" (turbulence in the ordinary sense), "pressure mode" (random sound), and "entropy mode" (temperature spottiness).[1]

The hot-wire anemometer can be used to measure turbulent fluctuations in the compressible flow regime;[1,2] however, its spatial resolution is greatly impaired by the limitation in frequency response that is aggravated by the high mean velocity of eddies with respect to the observer.

Shadow pictures of projectiles flying at supersonic velocities (Fig. 1) obtained by short duration spark sources (e.g., reference 3) exhibit turbulent regions with small scale details. The scale of the details is much finer than that which the hot-wire technique can resolve. Closer analysis reveals that the shadow picture emphasizes gradients because the technique responds to second spatial derivatives of the density fluctuations. The first obstacle to evaluating shadow pictures was the fact that the information is integrated along the light beam; therefore no recovery of point values is possible. When the objective of the measurement, however, is only the determination of averages, this integration effect is no great handicap.

It was shown that, with the assumptions of homogeneity and isotropy, the three dimensional spectrum of the turbulent density fluctuations is a Fourier-Bessel transform of the measured correlation function of the shadow picture.[4]

The more general mathematical problem of recovering the statistical properties (correlation or spectrum) of random fields from measurements obtained with probes or with other averaging methods has been extensively treated by the present authors.[5,6]

In the next section the method of evaluation will be outlined, and later two concrete examples will be shown.

II. TECHNIQUE AND ANALYSIS

Let $\rho(x_1,x_2,x_3)$ be the random density field statistically homogeneous and infinite in extent, where x_1, x_2, x_3 are Cartesian coordinates. A portion of the field, lying between x_3-l and x_3+l and extending to infinity in the other two directions, is removed from the rest of the field. Parallel light after passing through this slab of the field is incident on a photographic plate (Fig. 2). The local light flux arriving at the photographic plate will be more or less intense according to the distortion produced by the density fluctuations acting as concave or convex lenses. If the photographic plate is placed closer than the first focal point, there will be no singular

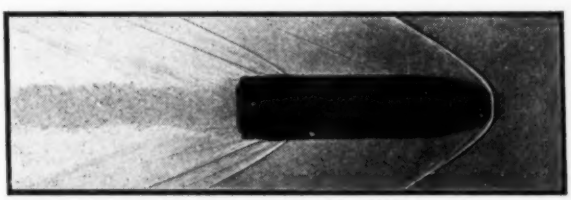


FIG. 1. Shadow picture of 19-mm shell in supersonic flow (courtesy of Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland).

* Present address: Department of Aeronautical Engineering, University of Michigan, Ann Arbor, Michigan.

[1] L. S. G. Kovasznay, J. Aeronaut. Sci. **20**, 657 (1953).
[2] L. S. G. Kovasznay, J. Aeronaut. Sci. **17**, 565 (1950).
[3] L. S. G. Kovasznay, Rev. Sci. Instr. **20**, 696 (1949).
[4] L. S. G. Kovasznay, *Heat Transfer and Fluid Mechanics Institute* (American Society of Mechanical Engineers, Berkeley, 1949).
[5] M. S. Uberoi, and L. S. G. Kovasznay, Project SQUID, Technical Report No. 30 (1951).
[6] M. S. Uberoi and L. S. G. Kovasznay, Quart. Appl. Math. **10**, 375 (1953).

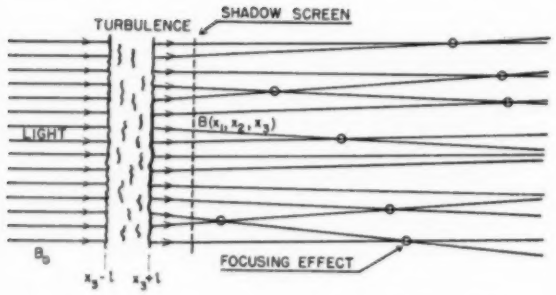


FIG. 2. Optical arrangement for taking shadow pictures.

points on the plate. The shadowgraph thus obtained still depends on the choice of position of the slab center along the x_3 axis. A single shadowgraph gives the mapped field for a fixed value of x_3, and the complete mapping of the density field consists of a continuous set of shadowgraphs, one for each value of x_3.

For small fluctuations of density the refractive index is a linear function of density, and under this assumption the response equation[7] becomes for a fixed x_3

$$h(x_1,x_2)=\frac{B_0-B(x_1,x_2)}{B(x_1,x_2)}$$
$$=\frac{LK}{\rho_0}\int_{x_3-l}^{x_3+l}\left(\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_2^2}\right)\rho(x_1,x_2,s)ds, \quad (1)$$

where B_0 is the intensity of incident light and B is the intensity of transmitted light arriving at the plate through a slab of random density field of thickness $2l$. L is the distance from the slab center to the photographic plate. ρ_0 is the mean density, and K is the Gladstone-Dale constant. Our problem is to determine the statistical properties (correlation or spectrum) of $\rho(x_1,x_2,x_3)$ from those of $h(x_1,x_2)$. The values of h can be determined by making a positive slide of the original shadowgraph and measuring its transparency ($\sim B/B_0$ for a linear photographic process) for each point x_1, x_2. The correlation or the spectrum of h can be computed from this data. The photographic plate (positive) has a transmittance

$$t(x_1,x_2)=t_0+\Delta t(x_1,x_2),$$

where t_0 is the background transparency and Δt is the fluctuating transparency whose average is zero and C is a constant depending on the photographic process. For small variation of light intensity and for linear photographic process

$$h(x_1,x_2)\cong 1-[B(x_1,x_2)/B_0]=-C\Delta t(x_1,x_2). \quad (2)$$

The correlation h can be determined[4] by making two identical positives of the original photograph and measuring the combined transparency of the two identical positives placed back to face and shifted from the matched position by distances ξ_1 and ξ_2. The average combined transparency can be determined by measuring it over a sufficiently large overlapping area. Thus the correlation is defined as

$$\beta(\xi_1,\xi_2)=\langle h(x_1,x_2)h(x_1+\xi_1,x_2+\xi_2)\rangle_{Av}$$
$$=C^2\langle\Delta t(x_1,x_2)\Delta t(x_1+\xi_1,x_2+\xi_2)\rangle_{Av} \quad (3)$$
$$=C^2[\langle t(x_1,x_2)t(x_1+\xi_1,x_2+\xi_2)\rangle_{Av}-t_0^2],$$

where we are taking space averages.

β is the correlation of h and, since h is statistically homogeneous, β is independent of x_1 and x_2. This rather simple method for determining β can be used even when B/B_0 is not small. In this case it is necessary to incorporate an appropriate nonlinear law into the photographic process in such a way that the fluctuating transparency of the positive is proportional to $(B_0-B)/B$ instead of B/B_0.

Next, we relate β, the correlation of the random photographic pattern h, to T, the correlation of the density field ρ. Substituting for h in the definition of β we have

$$\beta(\xi_1,\xi_2)=\langle h(x_1,x_2)h(x_1+\xi_1,x_2+\xi_2)\rangle_{Av}$$
$$=\left\langle\left(\frac{KL}{\rho_0}\right)^2\int_{x_3-l}^{x_3+l}\int_{x_3-l}^{x_3+l}\left(\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_2^2}\right)\right.$$
$$\times\left(\frac{\partial^2}{\partial x_1'^2}+\frac{\partial^2}{\partial x_2'^2}\right)\rho(x_1,x_2,s)$$
$$\left.\times\rho(x_1',x_2',\tau)ds d\tau\right\rangle, \quad (4)$$

where $x_1'=x_1+\xi_1$ and $x_2'=x_2+\xi_2$. Interchanging the order of averaging and of other operations and making use of the fact that the averages are functions of ξ_1, ξ_2, and $\tau-s$, we have

$$\beta(\xi_1,\xi_2)=\left(\frac{KL}{\rho_0}\right)^2\int_{x_3-l}^{x_3+l}ds\int_{x_3-l}^{x_3+l}d\tau$$
$$\times\left(\frac{\partial^2}{\partial\xi_1^2}+\frac{\partial^2}{\partial\xi_2^2}\right)^2 T(\xi_1,\xi_2,\tau-s), \quad (5)$$

where $T(\xi_1,\xi_2,\xi_3)$ is the correlation of the density field $\rho(x_1,x_2,x_3)$. Let $y=\tau-s$ and s be the new variables. Integrating first with respect to s and making use of the fact that for statistically homogeneous and isotropic density fluctuations β and τ are functions of $r=(\xi_1^2+\xi_2^2)^{\frac{1}{2}}$ and $(r^2+y^2)^{\frac{1}{2}}$ respectively, we have

$$\beta(r)=2\left(\frac{KL}{\rho_0}\right)^2\int_0^{2l}(2l-y)\left(\frac{\partial^2}{\partial r_2}+\frac{1}{r}\frac{\partial}{\partial r}\right)$$
$$\times\left(\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}\right)T(r^2+y^2)^{\frac{1}{2}}dy. \quad (6)$$

An interesting situation arises when the thickness of the

[7] F. J. Weyl, NAVORD Report No. 211-45 (1945).

slab is much larger than any turbulence scale. In this case we may consider the slab thickness as effectively infinite ($l\to\infty$):

$$\beta(r)\cong 4l\left(\frac{KL}{\rho_0}\right)^2\int_0^{2l\to\infty}\left(\frac{\partial^2}{\partial r^2}+\frac{1}{r}\frac{\partial}{\partial r}\right)^2 T(r^2+y^2)^{\frac{1}{2}}dy. \quad (7)$$

As l becomes larger the discrepancy between Eqs. (6) and (7) becomes smaller. The solution of Eq. (7) is[5,6]

$$T(r)=\frac{1}{8\pi l}\left(\frac{\rho_0}{KL}\right)^2\int_r^\infty \beta(y) \times\left[y\frac{(2r^2+y^2)}{r}\cos^{-1}\left|\frac{r}{y}\right|-3y(y^2-r^2)^{\frac{1}{2}}\right]dy. \quad (8)$$

We see that the correlation of the three dimensional random density field can be recovered from the correlation of the two dimensional shadow picture. It was shown in reference (4) that the three dimensional spectrum of the density field $\rho(x)$ can be recovered from the measured correlation of the shadow picture. The statistical representation of a random field by spectrum is equivalent to its representation by correlation. In the present investigation we will mainly consider the correlation.

Expanding Eq. (8) for small values of r and making use of the fact that $T(r)$ tends to a finite value for small r, we have

$$\int_0^\infty \beta(\xi)\xi d\xi\equiv\int_0^\infty \beta(\xi)\xi^3 d\xi\equiv 0. \quad (9)$$

For the isotropic case and infinite light path, $\beta(r)$ must always oscillate so as to satisfy Eq. (9). In Fig. 3 we have compared $\beta(r)$ with $T(r)$ for the special case when $T(r)$ is Gaussian.

III. EXPERIMENT I—WAKE OF PROJECTILE

The technique for measurement of turbulent density fluctuations was first applied to the wake of a 19-mm shell flying at a velocity of approximately 1700 ft/sec in atmospheric air (Fig. 1). The simplest optical arrangement was used. The picture was taken with a high-intensity, short-duration (few microseconds) spark without collimating lens. This procedure gives a slight magnification, but its simplicity and especially the absence of any refracting or reflecting medium (collimating lenses or mirrors) gives a very high contrast.

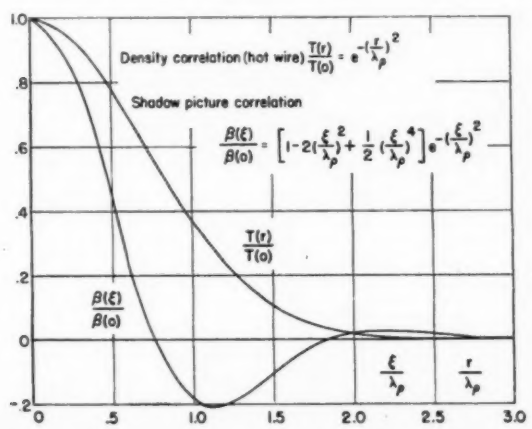


FIG. 3. Shadow correlation for Gaussian density correlation.

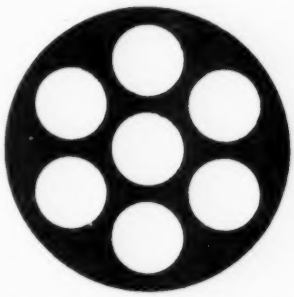


FIG. 4. "Swiss-cheese" grid projectile; shadow picture of "Swiss-Cheese" projectile in supersonic flow (courtesy of Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland).

It was a surprise to find that the visible turbulent wake pattern persists quite a long distance behind the projectile. It is believed that the violent shearing in the wake just behind the projectile produces high entropy gradients with the result that the flow becomes "entropy spotty" due to later turbulent mixing. As the turbulent motion decays (because of not only viscosity but also sound radiation), the remaining effect is a temperature and therefore density spottiness that decays only by the action of molecular heat conduction.

The wake of ordinary shells was used in preliminary measurement, but the turbulence in the wake is hardly homogeneous and isotropic. The best shadowgraphs are obtained for turbulent flows with free boundaries. Homogeneous and isotropic turbulence can be produced by placing a grid in the working section of a wind tunnel; however, high speed tunnels are expensive to operate and none was available for this work at that time. The wind tunnel has another disadvantage: the density fluctuations in the turbulent boundary layers along the tunnel wall are large, and as the light goes through both boundary layers the additional shadow effects become large and may even overshadow the homogeneous density field in the free stream.

It was decided to produce an approximately homogeneous and isotropic turbulent field of fairly large extent and free boundaries. This was accomplished by the use of a "grid projectile" (Fig. 4). (See also cover.) The total diameter was 3.7 cm and the thickness was 0.8 cm. Seven holes formed half of the total area. The firing of such an unconventional projectile required the use of a sabot. Flight of the "swiss-cheese" was quite satisfactory, and Fig. 4 shows the flow at Mach number 1.03. The jets from the individual holes mix and an approximate normal shock region develops. The tur-

bulence farther behind seems to be reasonably homogeneous and isotropic.

The measured correlation of the shadow picture is shown in Fig. 5(a). No absolute calibration of the photographic plate was possible, so that actual measurement can only give the correlation coefficient $\beta(r)/\beta(0)$. The location of the horizontal axis, $\beta(r)/\beta(0)=0$, is determined from the asymptotic value of $\beta(r)/\beta(0)$ for large values of r. At the time of the original measurement, it was believed that the shape of $\beta(r)/\beta(0)$ was most significant, and, as its measurements for large values of r were not made, the position of the horizontal axis is somewhat uncertain. The measured curve was completed with an assumed portion to secure zero first and third moments as required by Eq. (9). Figure 5(b) shows the computed density correlation coefficient.

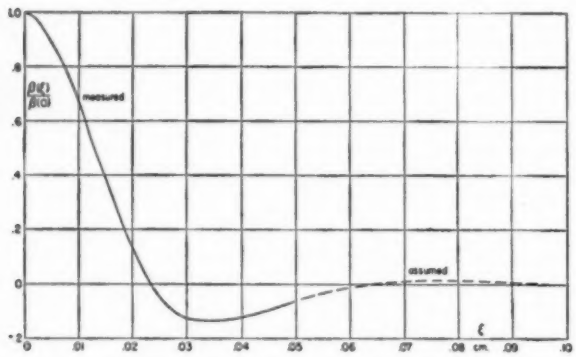


FIG. 5a. Shadow correlation for compressible turbulence.

Often the sole purpose of correlation measurement is the determination of scales of turbulence. In such a case it is not necessary to compute $T(r)$ since L_ρ and λ_ρ can be expressed directly in terms of $\beta(r)$ from Eq. (8).

$$L_\rho=\int_0^\infty T(r)dr/T(0)$$

$$=-\frac{\pi}{8}\int_0^\infty \beta(\xi)\xi^3\log\xi d\xi \Big/ \int_0^\infty \beta(\xi)\xi^2 d\xi. \quad (10)$$

$$\lambda_\rho{}^2=-2T(0)T''(0)$$

$$=-6\int_0^\infty \beta(\xi)\xi^2 d\xi \Big/ \int_0^\infty \beta(\xi)d\xi. \quad (11)$$

In the turbulent wake produced by the grid projectile, we found $L_\rho=0.0295$ cm and $\lambda_\rho=0.030$ cm. The estimated equivalent "mesh length" M of the grid projectile was about 1 cm and the mean velocity U was of the order of sound velocity. The grid Reynolds number

$$R_M=MU/\nu\cong 2\times 10^5,$$

where ν is the kinematic viscosity. There is no information, however, on the turbulence level. For merely estimating whether the data obtained is reasonable or not, let us assume that $u'/U=0.10$ (u' is the rms turbulent velocity) and that the relation between λ (the microscale of turbulent velocity field) and λ_ρ is the theoretical value found by Corrsin.[8] For large Reynolds number he found that

$$2\lambda^2/\lambda_\rho{}^2=\text{Prandtl number}=0.7 \text{ for air.}$$

Thus $\lambda=0.6\lambda_\rho$, or in this case $\lambda=0.018$ cm.

The turbulence may be characterized by the Reynolds number, and with the above assumed turbulence level

$$R_\lambda=\frac{u'}{U}\frac{\lambda}{M}R_M=360.$$

The smallness of the microscale compared with the scale of the turbulence producing mechanism is a direct consequence of the high value of R_λ. Such high values of R_λ are not encountered usually in low-speed tunnels behind the grids.

The smallness of L_ρ appeared rather suspicious, since, for high Reynolds number turbulence, the integral scale L_ρ is usually much larger than the microscale λ_ρ. However, in our case it was found that $\lambda_\rho/L_\rho\cong 1$. It is quite clear from Eq. (10) that L_ρ decisively depends on the behavior of $\beta(r)$ for larger r values. However, in this region the curve was merely guessed in accordance with requirements expressed in Eq. (9). Knowing that the shadow method is sensitive only to the second spatial derivatives indicates that it de-emphasizes the low wave number contribution that determines L_ρ. The microscale λ_ρ, on the other hand, depends more on the

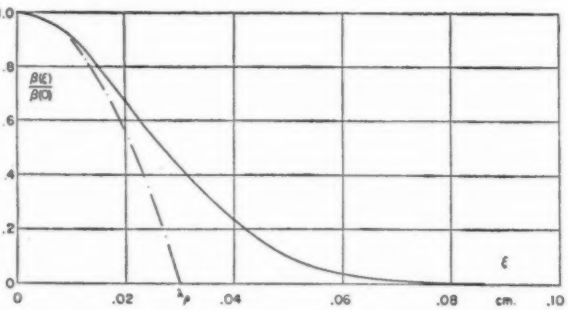


FIG. 5b. Density correlation for compressible turbulence.

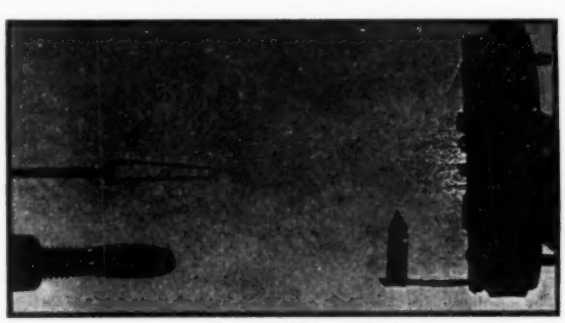


FIG. 6. Shadow picture of heated jet and hot-wire probe.

[8] S. Corrsin, J. Aeronaut. Sci. 18, 417 (1951).

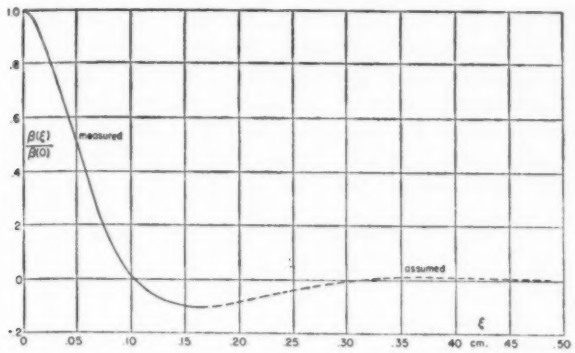


FIG. 7a. Shadow correlation for incompressible turbulence.

higher wave numbers and therefore is expected to be more reliable.

IV. EXPERIMENT II—HEATED LOW SPEED JET

The results obtained from the shadowgraph of approximately homogeneous and isotropic turbulence produced by a grid projectile looked promising enough to warrant further experimentation. However, it is not possible to check these results with another independent measurement. In the low Mach number case, however, the changes in pressure due to the motion of the fluid are small, and the density can be regarded as inversely proportional to the absolute temperature. For small changes, therefore, there is a linear relation between density and temperature. At low speeds the hot wire can be used as a resistance thermometer to measure temperature fluctuations.[9]

It was decided to produce an approximately homogeneous and isotropic low-speed turbulent flow with density fluctuations and with free boundaries.[10] This was accomplished by placing an electrically heated grid at the orifice of a small nozzle. The orifice was square (2.5×2.5 cm) discharging into free air at a velocity of 1080 cm/sec. The grid was made of 0.080 cm diameter nichrome wires placed 0.62 cm apart. Figure 6 shows the shadowgraph picture of the flow including the shadow of a hot-wire probe. Most measurements were made 2 cm downstream from the grid where the pointer is located in the picture. Figure 7(a) shows the measured correlation of the shadow plates, and Fig. 7(b) shows the computed density correlation coefficient. Again it was necesary to add an assumed portion to the measured curve in order to make its first and third moments zero. The microscale and integral scale values were found

$$\lambda_\rho = 0.145 \text{ cm} \quad \text{and} \quad L_\rho = 0.138 \text{ cm}.$$

In order to check the validity of these values, independent measurement of temperature (~density) fluctuations was made with a hot-wire anemometer by using it as a resistance thermometer. The statistical homogeneity of the density fluctuations was checked by measuring $E_1(k_1)$ the one-dimensional spectrum of density fluctuations 2 cm downstream from the grid and at the center-line of the nozzle, and also at two other positions 0.5 cm on either side of the center-line. Figure 8 shows that there is good agreement among the three measurements. In order to compare hot-wire measurements with shadow measurement, we can compute λ_ρ and L_ρ from the spectrum $E_1(k_1)$.†

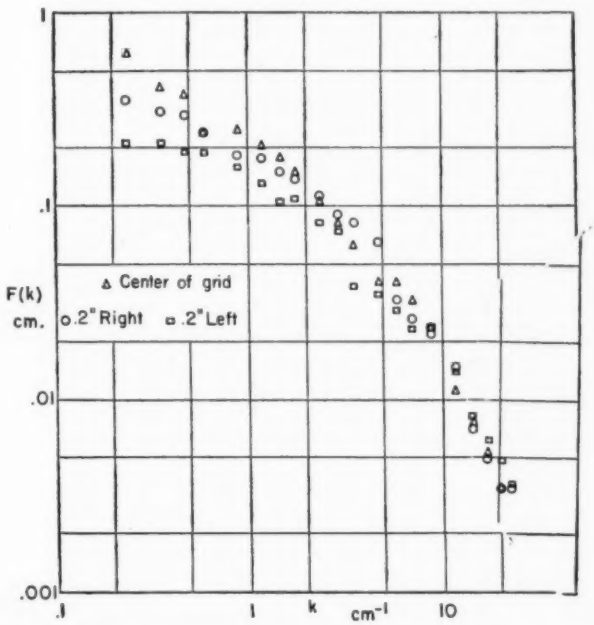


FIG. 8. Energy spectrum of temperature (density) fluctuation measured by hot wire.

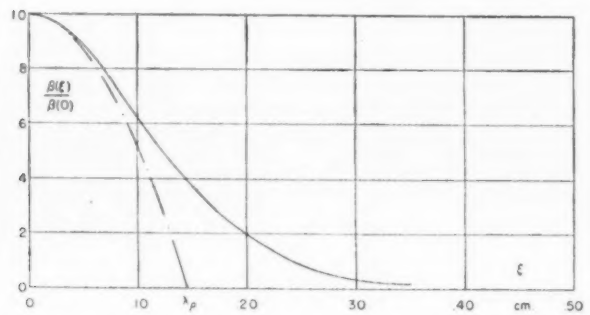


FIG. 7b. Density correlation for incompressible turbulence.

$$T(r)/T(0) = \int_0^\infty E_1(k_1) \cos k_1 r \, dk_1;$$

$$E_1(k_1) = \frac{2}{\pi T(0)} \int_0^\infty T(r) \cos k_1 r \, dr.$$

[9] S. Corrsin, Rev. Sci. Instr. **18**, (1947).

[10] L. S. G. Kovasznay and P. C. Clarken, Project SQUID, Technical Report No. 42 (1953).

† In reference 10 $E_1(k_1)$ is given as a function of $n = k_1/2\pi U$.

It follows from the definition of λ_ρ and L_ρ (see Eqs. (10) and (11)) that

$$\lambda_\rho=\int_0^\infty E_1(k_1)k_1^2dk \quad \text{and} \quad L_\rho=\frac{\pi}{2}E_1(0).$$

The values of λ_ρ and L_ρ computed from the spectrum were found to be $\lambda_\rho=0.123$ cm and $L_\rho=0.850$ cm. The agreement for the microscale is reasonable but for the integral scale is very poor.

V. CONCLUSIONS

The shadow picture carries the necessary information to determine the correlation function of the turbulent density fluctuations. However, the method strongly emphasizes the high wave number contributions. The microscale that depends largely on the medium and the higher wave number portion of the spectrum is therefore better recovered than the integral scale. The shape of the density correlation curve cannot be recovered too well unless the shadow correlation is determined with great accuracy, especially its asymptotic behavior for large distances.

It appears that the shadow method may be used to contribute additional information on the behavior of the smallest eddies in cases where the behavior of larger eddies is already well-known, or in cases where no other measurements are possible and an estimate microscale is necessary.

VI. ACKNOWLEDGMENT

The present work was carried out under the sponsorship of SQUID Project Contract N6ori-105 Task Order III and the work of one of the authors (M.S.U.) was also supported by the Office of Naval Research under Contract Nonr-248(38). The authors are indebted to Miss Patricia C. Clarken for her assistance in the project.

The shadow picture on the cover and Figs. 1 and 4 were taken in the firing range of the Ballistic Research Laboratory, Aberdeen Proving Ground, Maryland, and the authors acknowledge the enthusiastic support of the staff.

JOURNAL OF APPLIED PHYSICS VOLUME 26, NUMBER 1 JANUARY, 1955

Analysis of the Effect of Various Factors on Metal Transfer and Wear between Specimen Pairs of Same Metal and Same Shape. I. The Basic Scheme of Formulation of Metal Transfer and Wear*

I-Ming Feng
Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received May 18, 1953; revised manuscript received May 22, 1954)

The fundamental scheme of formulation of the amount of metal transfer and wear is based upon the theory suggested in the previous paper [J. Appl. Phys. **23**, 1011–1019 (1953)]. The basic terms used in formulating metal transfer and wear are (1) the potential amount of metal transfer M_t^*; (2) the mechanical factor of the chance of metal transfer ϕ_m; (3) the thermal factor of the chance of metal transfer ψ_θ; and (4) the adhesion factor of the chance of metal transfer X_a. The expression for the amount of metal transfer M_t from one specimen to another is $M_t=M_t^*\phi_m\psi_\theta$, when ψ_θ is bigger than X_a, and is $M_t=M_t^*\phi_m X_a$ when X_a is bigger than ψ_θ. The equation for the amount of wear M_w represented by the weight loss of a specimen is $M_w=M_t^*\phi_m(1-\psi_\theta)$, when ψ_θ is bigger than X_a, and is $M_w=M_t^*\phi_m(1-X_a)$, when X_a is bigger than ψ_θ.

INTRODUCTION

METAL transfer and wear impair the rubbing parts by changing their weight, dimensions, shape, surface condition, etc. In the following discussion, weight change will be discussed and formulated. The reason is as follows: Weight change is admittedly only a measure of one phase of the damage produced by wear and, therefore, is by no means a perfect description of metal transfer or wear. As compared with other changes produced by metal transfer and wear, however, the change in weight is the only one which can usually be measured quantitatively with satisfactory accuracy. Dimensional change is usually more difficult to determine with reproducible and accurate results. Other changes can only be described qualitatively. The best description would be a combination of quantitative data of weight change and at the same time qualitative pictures of other changes. In the absence of one single perfect measure in describing metal transfer or wear, the weight change will be discussed and formulated as a function of various factors in this series of publications.

* This is the first paper in a series of publications that contain a part of the dissertation submitted by the author in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of Michigan.

METAL TRANSFER

According to the mechanism of metal transfer suggested in the previous paper,[1] the amount of metal transfer M_t from one specimen to the other is expected to be the product of the following three factors:

(1) The potential amount of metal transfer M_t^*.
(2) The mechanical factor of the chance of metal transfer ϕ_m.
(3) The thermal factor of the chance of metal transfer ψ_θ or the adhesion factor of the chance of metal transfer X_a, depending upon which one is bigger.

The potential amount of metal transfer M_t^* is defined as the amount of metal that would be transferred if, at every plastically roughened contacting high spot, a blob of metal were sheared off and attached to the opposite surface. M_t^* is assumed to be proportional to the area A_p of the plastically deformed region or regions, and the average depth d_p of the strain-hardened region or regions. The dependence of M_t^* on the distance traveled during relative motion is not linear and is, therefore, given by a factor D.† Thus we have

$$M_t^* = c_m \rho A_p d_p D, \tag{1.01}$$

where ρ is the density, and c_m is the constant of proportionality. The total actually contacting area A consists of two parts: the elastically deformed region, or regions, of total area A_e, and the plastically deformed region, or regions, of total area A_p. The area A_p instead of A is involved in Eq. (1.01), because the elastically deformed region that is not mechanically interlocked by roughening contributes little to the amount of metal transfer.

The primary cause of metal transfer is the mechanical interlocking of the roughened interface. A pair of weakly interlocked high spots may separate again along the original interface and thus contribute nothing to the metal transfer. The greater the degree of roughening of the interface, the higher will be the resistance to separation along the original interface. The mechanical factor ϕ_m of the chance of metal transfer is assigned to take care of the dependence of the chance of metal transfer upon the firmness of the mechanical interlocking.

If metal is to be transferred from one specimen to another, not only must breakage occur inside the bulk rather than along the original interface, but also some interaction must operate to hold the sheared-off peak or peaks in their new places. The interaction can be either welding resulting from the diffusion during the temperature flash, as explained in reference 1, or the adhesive force. It is also possible that both of them are operating.

In ordinary atmosphere, metallic surfaces are always contaminated by a layer of surface film which prevents metallic adhesion. As has been discussed in the previous paper,[1] metal-to-metal contact could still be established to a certain extent by the roughening of the interface under the special condition that the surface film is brittle and much thinner than the average depth of roughening of the interface (see Fig. 2b in reference 1). For a certain thickness of surface film, the chance to establish some metal-to-metal contact inside the contacting area or areas will be larger for a greater average depth of roughening of the interface. Therefore, the need is obvious for having an adhesion factor of the chance of metal transfer X_a which is defined as the chance that metal-to-metal contact established during plastic roughening of the interface will be strong enough to attach the sheared-off peak to the high spot on the opponent surface.

If a peak is sheared off from a high spot and no adhesion is present, the sheared-off blob cannot become a piece of transferred metal unless the diffusion process during the temperature flash is rapid enough to produce welding. The more rapid the diffusion process is, the greater the chance for the blob of metal sheared off from a high spot to be welded to another high spot on the opposite surface. The thermal factor of the chance of metal transfer ψ_θ is provided for this purpose. It is defined as the chance that the temperature flash will weld the sheared-off blob to the opposite surface.

Thus, the basic idea involved in formulating the amount of metal transfer from one specimen to its opponent specimen without taking resoftening and back transfer into consideration can be expressed as follows:

$$M_t = M_t^* \phi_m \psi_\theta, \quad \text{when} \quad \psi_\theta > X_a \tag{1.02a}$$

and

$$M_t = M_t^* \phi_m X_a, \quad \text{when} \quad X_a > \psi_\theta. \tag{1.02b}$$

The use of Eq. (1.02a) or (1.02b) is determined by which one of these two factors is operating, ψ_θ or X_a. When a large number of high spots on two surfaces are contacting, it is usually the case that, for all contacting high spots, the same one of these two factors ψ_θ and X_a is larger. The reason is as follows: The question of which one of these two factors is larger is mainly determined by the surface film. Except under special conditions, it is a reasonable assumption that all high spots on the same surface have the same surface condition, and hence the same factor is operating for all high spots. Therefore, Eqs. (1.02a) and (1.02b) can be applied to all cases regardless of the number of high spots in actual contact.

Resoftening, Back Transfer, and Net Amount of Metal Transfer

The phenomenon of metal transfer is further complicated in a great degree by the resoftening of the strain-hardened transferred blob of metal and back transfer of transferred metal. The strain-hardened blob receives heat treatment during the temperature rise,

[1] I. M. Feng, J. Appl. Phys. **23**, 1011–1019 (1952).
† Detailed discussion about D will be given in a later publication.

resoftens because of the heat of temperature flash, and stays soft after the temperature flash is over. The resoftened blob may be transferred back to its mother specimen either partly or completely. The chance for the transferred blob to be transferred back to its mother specimen is largely dependent on the degree of resoftening, which in turn depends upon the temperature flash.

Let $(M_t)_{ab}$ denote the amount of A atoms transferred from A to B. Further, let $(M_t)_{bab}$ denote the amount of back transfer of B atoms from A to their mother specimen B. The total amount of metal transfer from A to B is then given by

$$M_t=(M_t)_{ab}+(M_t)_{bab}, \tag{1.03}$$

or

$$(M_t)_{ab}=M_t-(M_t)_{bab}. \tag{1.03$_a$}$$

The net amount of A atoms transferred from A to B, $(M_t)_{ab,\,net}$, is given by $(M_t)_{ab}$ minus the amount of A atoms back-transferred from B, denoted by $(M_t)_{aba}$.

$$\begin{aligned}(M_t)_{ab,\,net}&=(M_t)_{ab}-(M_t)_{aba}\\&=M_t-(M_t)_{bab}-(M_t)_{aba}.\end{aligned} \tag{1.04}$$

Because specimens A and B are made of identical metal, and have same shape,

$$(M_t)_{aba}=(M_t)_{bab}. \tag{1.05}$$

Thus, Eq. (1.04) becomes

$$(M_t)_{ab,\,net}=M_t-2(M_t)_{aba}. \tag{1.04$_a$}$$

Differentiating M_t once with respect to the distance of travel L, Eq. (1.03) becomes

$$\partial M_t/\partial L=\partial(M_t)_{ab}/\partial L+\partial(M_t)_{bab}/\partial L.$$

At the very beginning,

$$\partial/\partial L\ (M_t)_{bab}=0,$$

and

$$\partial/\partial L\ M_t=\partial/\partial L\ (M_t)_{ab}.$$

As the distance of travel increases, $\partial(M_t)_{bab}/\partial L$ varies. Let, at any instant, $\partial(M_t)_{bab}/\partial L$ be a fraction Λ of $\partial M_t/\partial L$, namely,

$$\partial(M_t)_{bab}/\partial L=\Lambda\partial M_t/\partial L. \tag{1.06}$$

Here, Λ is influenced by resoftening as well as by the distance of travel. From Eq. (1.06), it follows,

$$(M_t)_{ab,\,net}=M_t-2\int_0^L\Lambda(\partial M_t/\partial L)dL, \tag{1.07}$$

where M_t is given by Eq. (1.02$_a$) and (1.02$_b$) in terms of M_t^*, ϕ_m, X_a, and ψ_θ.

Wear

Contributing positive results to the wear of the specimen A represented by its weight loss M_w are the following terms:

(1) The amount of loose wear particles worn away from the specimen A, denoted by $(M_w')_a$.
(2) The amount of A metal transferred from A to B.
(3) The amount of B metal back-transferred from A to B.

The terms which contribute negative results to M_w are as follows:

(4) The amount of B metal transferred to A.
(5) The amount of A metal back-transferred from B to A.

Since the specimens A and B under consideration have practically identical properties, the amount of metal transfer from A to B is compensated by the amount of metal transfer from B to A, so that the last four terms do not contribute any net result to the weight loss of the specimen A. Although what actually happens is quite complicated, the amount of wear represented by weight loss of the specimen is simply equal to the amount of the loose wear particles worn away from the specimens in the case of two rubbing specimens of the same metal.

$$M_w=(M_w')_a. \tag{1.08}$$

This result indicates that the weight loss is caused by wear only and is not perturbed by primary transfer, nor by back transfer.

When two specimens rub together, metal transfer and wear take place simultaneously. The product $M_t^*\phi_m$ represents the fraction of the potential amount of metal transfer M_t^*, which will leave the specimen to become either loose wear particles or transferred metal. The portion of $M_t^*\phi_m$ that is transferred from specimen A to specimen B is represented by M_t, and the remaining part which is not transferred leaves the specimen A as loose wear particles, represented by $(M_w')_a$ that is equal to M_w. In other words,

$$M_t^*\phi_m=M_t+M_w. \tag{1.09}$$

Thus, the basic idea involved in formulating the amount of wear can be expressed as follows:

$$M_w=M_t^*\phi_m(1-\psi_\theta),\quad \text{when}\quad \psi_\theta>X_a, \tag{1.10$_a$}$$

and

$$M_w=M_t^*\phi_m(1-X_a),\quad \text{when}\quad X_a>\psi_\theta. \tag{1.10$_b$}$$

DISCUSSION

In formulating metal transfer and wear between specimen pairs of the same metal and the same shape, the amount of metal transfer, back transfer, wear, etc., is theoretically the same for both specimens. This finds its support from Kehl and Siebel's results[2] of wear experiments. In their experiments, two specimens of the same ring shape were made of the same metal and were arranged side by side with the side faces of the rings in contact. Their results indicate that the weight loss of

[2] B. Kehl and E. Siebel, Arch. Eisenhüttenw. 9, 563–570 (1936).

the stationary specimen was practically the same as that of the moving specimen.

The equations for metal transfer and wear are expressed as a function of such terms as M_t^*, ϕ_m, ψ_θ, X_a, and Λ. Each of these terms is a function of various factors, such as load, speed, distance of travel, etc.

The terms ϕ_m, ψ_θ, and X_a can have any value from zero up to one. Under the conditions that X_a and ψ_θ reach their upper limiting values, the amount of wear M_w is nil, according to Eqs. (1.10_a) and (1.10_b). This is a very interesting case indeed. The weight loss vanishes under such conditions because every sheared-off peak will stick to its opponent surface and give no loose wear particle.

If the term $\partial M_t/\partial L$ in Eq. (1.07) is assumed to be constant, then

$$(M_t)_{ab,\,net}=M_t\left(1-\frac{2}{L}\int_0^L \Lambda dL\right). \qquad (1.11)$$

When the conditions are such that the transferred blob of metal is completely resoftened by the temperature flash, Λ becomes an increasing function of L alone. In such a case, $\partial(M_t)_{ab,\,net}/\partial L$ approaches zero at a very large distance of travel L. In other words, when L becomes very large, $\partial M_t/\partial L$ for transfer from A to B is evenly divided between $\partial(M_t)_{ab}/\partial L$ and $\partial(M_t)_{bab}/\partial L$; and $\partial M_t/\partial L$ for transfer from B to A is also evenly divided between $\partial(M_t)_{ba}/\partial L$ and $\partial(M_t)_{aba}/\partial L$. That is,

$$\partial(M_t)_{bab}/\partial L=\partial(M_t)_{aba}/\partial L=\tfrac{1}{2}\,\partial M_t/\partial L.$$

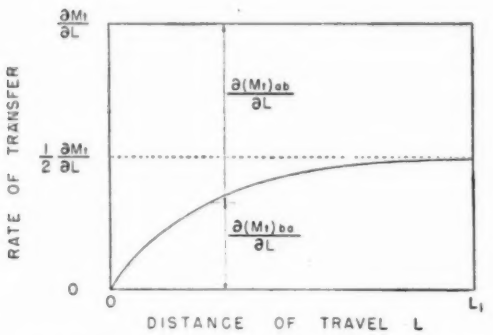


FIG. 1. Sketch illustrating the variation of $\partial(M_t)_{ba}/\partial L$ as an increasing function of L. $\partial M_t/\partial L$ is assumed to be constant.

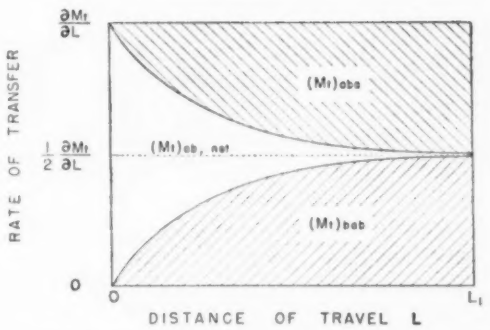


FIG. 2. Sketch showing the variation of $(M_t)_{ab,\,net}$ that is represented by the unshaded area.

Assume that

$$\Lambda=\tfrac{1}{2}(1-e^{-L/l}), \qquad (1.12)$$

which is an increasing function starting from zero at $L=0$ and approaching $\frac{1}{2}$ asymptotically at large L as shown in Fig. 1. The constant e has the same dimension as L. The quantity $(M_t)_{aba}=(M_t)_{bab}=(M_t/L)\int_0^L \Lambda dL$ is represented by the shaded area in Fig. 2. Using Eq. (1.04_a),

$$(M_t)_{ab,\,net}=\frac{M_t^*}{L}\phi_m\psi_{el}(1-e^{-L/l}), \qquad (1.11_a)$$

that is the unshaded area as illustrated in Fig. 2.

When the transferred blob of metal is not completely softened by the temperature flash, the dependence of Λ on resoftening and the distance of travel is unfortunately exceedingly complicated. This complication makes it impractical to express $(M_t)_{ab,\,net}$ in terms of load, speed, etc., and offsets the advantage that $(M_t)_{ab,\,net}$ can easily be determined experimentally by radioactive-tracer technique. In fact, M_t actually serves the purpose of describing the damage caused by metal transfer better than $(M_t)_{ab,\,net}$. For example, in the preceding case, metal transfer continues to cause a great deal of damage even at very large distance of travel when the quantity $\partial(M_t)_{ab,\,net}/\partial L$ is almost zero. Therefore, the amount M_t or the rate $\partial M_t/\partial L$ is actually a better measure of damage caused by metal transfer and will be discussed in analyzing the effect of various factors on metal transfer. It also has the obvious advantage that the sum of M_t and M_w is equal to $M_t^*\phi_m$ as given by Eq. (1.09).

Analysis of the Effect of Various Factors on Metal Transfer and Wear between Specimen Pairs of Same Metal and Same Shape. II. Effect of the Surrounding Atmosphere

I-Ming Feng
Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received May 22, 1954)*

The discussion in this paper is limited to the case where the surface layer is thinner than the average depth of roughening. The surrounding atmosphere influences the adhesion factor X_a through its effect on adhesion that depends on such factors as the nature and thickness of the surface layer and also the degree of plastic roughening. The influence of the surface layer on the thermal factor ψ_θ is similar to its influence on X_a. Because of the lack of the quantitative knowledge about its effect on adhesion and diffusion, the effect of surrounding atmosphere on metal transfer and wear is discussed qualitatively. The dependence upon such factors as the temperature, gas pressure, time of exposure, etc., of the effect of physically adsorbed and chemical-compound surface layers on metal transfer and wear is also discussed in this paper.

INTRODUCTION

THE fundamental scheme of formulation of the amount of metal transfer and wear has been presented in a previous paper.[1] The formulation will now be applied to the analysis of the effect of various factors on metal transfer and wear. Unless otherwise stated, it will be understood that factors other than the one under consideration are all kept unchanged in the analysis.

ADHESION

In high vacuum, metallic adhesion between a contacting pair of perfectly clean high spots will join them together by continuous metallic bonds extending from one high spot to another. When exposed to a gaseous atmosphere a perfectly clean surface adsorbs gas molecules. The adsorption may be merely physical in nature, or it may be a combination of physical and activated adsorption. If the gas is active, a surface layer made of chemical compound will be formed. Among these three different kinds of surface layer, a physically adsorbed layer is expected to be least effective, and a chemical-compound layer is considered to be most effective in reducing the adhesion.

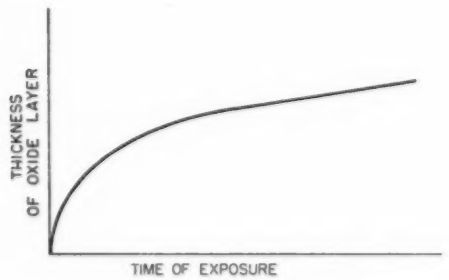


FIG. 1. Increase in thickness of oxide layer as a function of the time of exposure.

Bowden and Young[2] observed that an adsorbed-gas layer prevents complete seizure. Gwathmey, Leidheiser, and Smith[3] have performed experiments in measuring the resistance to normal separation between two contacting copper single crystals. They observed that a force of the order of, or slightly higher than, the normal load that was used to push the surfaces together was required to separate in the normal direction two oxide-free surfaces of contacting copper single crystals. Their experiments with similar copper single crystals which were covered by oxide surface film formed in air showed no detectable adhesion.

THICKNESS

Regarding the effect of the thickness of the surface layer, there are evidences which show that the thicker the layer, the more effective it is in reducing adhesion, provided that all other conditions are kept the same. The formation of a chemical compound is a relatively slow (slow as compared to the adsorption of gas by solid surface) rate-determining process. When all other conditions are kept unchanged, the thickness of the surface layer of chemical compound will increase with increasing time of exposure to the gas as shown in Fig. 1 (see also reference 4). McFarlane and Tabor[5] showed that the adhesion as measured by the resistance to normal separation of a clean steel ball on a clean lead surface decreases with increasing thickness of oxide layer on the metal surface as the result of prolonged exposure to the air. As the time of exposure increases, the rate of decrease of adhesion falls off because of the decreased rate of thickening of the oxide layer. The adhesion is reduced practically to zero after about eight hours of exposure. They also showed a similar result for

* The first, unrevised manuscript was received May 18, 1953.
[1] I-M. Feng, preceding paper, J. Appl. Phys. **26**, 24 (1955).
[2] F. P. Bowden and J. E. Young, Proc. Roy. Soc. (London) **A208**, 311–325 (1951).
[3] Gwathmey, Leidheiser, and Smith, National Advisory Comm. Aeronaut. Technical Note No. 1461.
[4] N. B. Pilling and R. E. Bedworth, J. Inst. Metals **29**, 529 (1923).
[5] J. S. McFarlane and D. Tabor, Proc. Roy. Soc (London) **A202**, 224 (1950).

the adhesion of a clean steel ball on a clean indium surface, except with a much lower rate of falling of the adhesion because of the much lower rate of oxidation of indium as compared with lead.

Numerous experimental results have shown that there is no detectable resistance to normal separation of two contacting oxide-covered surfaces The absence of resistance to normal separation is an indication that the adhesion between oxide-covered surfaces is at most a feeble one. Although these results of no normal-separation resistance have been attributed by some investigators[6,7] to the elastic recovery of the bulk material during unloading rather than the absence of adhesion, the elastic-recovery explanation is not satisfactory. For example, consider again the experimental result cited before regarding the adhesion of a clean steel ball on clean lead surface. The adhesion is practically zero after about eight hours of exposure of the lead surface to air. For a metal like lead with resoftening temperature below room temperature, the elastic recovery is negligibly small and, therefore, can hardly be the reason of no detectable resistance to normal separation.

PLASTIC ROUGHENING

Although the surface layer of oxide formed in atmospheric air is usually thick enough to prevent adhesion under moderate load as a result of the very rapid initial rate of thickening of the oxide layer, a certain amount of metal-to-metal contact could be established during the large plastic roughening of the interface caused by a very heavy load, as has been illustrated in the previous paper.[8] When other conditions are kept unchanged, the greater the degree of plastic roughening of the interface, the larger the amount of metal-to-metal contact and, hence, the stronger the total adhesion will be.

If the oxide-surface film is brittle and thin, excessive plastic roughening can cause enough metal-to-metal contact to join two surfaces together.† This technique of joining two carefully-cleaned metallic parts at room temperature employing extreme plastic roughening has been used commercially and is often called "cold-welding." However, the plastic roughening involved in rubbing metallic surfaces hardly reaches such a magnitude as that employed in the cold-welding technique.

[6] H. Ernst and M. E. Merchant, Proc. Special Summer Conferences on Friction and Surface Finish, M.I.T., Cambridge, Massachusetts, pp. 206–211.

[7] Bowden, Moore, and Tabor, J. Appl. Phys. **14**, 80–91 (1943).

[8] I-M. Feng, J. Appl. Phys. **23**, 1011 (1952).

† The same mechanism of plastic roughening operates in pressing metal powders to form coherent masses. A clean surface is essential for securing satisfactory results. For example, a reduction treatment for decreasing the thickness of the surface-oxide film on metal powder is usually involved in the drying operation of metal powders in powder metallurgy. In the drying operation, heating the metal powders also serves the purpose of annealing them, which gives greater plastic roughening under a given pressure during the pressing operation.

THIN SURFACE LAYER

In most cases, the surrounding atmosphere does not change the potential amount of metal transfer M_t^*. The effect of the thickness of the surface film on ϕ_m has already been discussed in reference 8. If the thickness of the surface film is small as compared to the average depth of plastic roughening, the surrounding atmosphere does not cause appreciable influence on the mechanical factor ϕ_m. Since the sum of M_t and M_w is equal to $M_t^*\phi_m$ (Eq. (1.09) of reference 1), the effect of surrounding atmosphere is merely changing the distribution of the total damage $M_t\phi_m$ between two different kinds of detriment, metal transfer M_t and wear M_w. Between them, wear is sometimes preferred because it at least does not prevent continuous operation. Thus, modification of the surrounding atmosphere provides means for controlling the form of damage.

The analysis given in this paper will be limited to the case that the surface layer is thinner than the average depth of plastic roughening. Under this condition, the surrounding atmosphere influences only ψ_θ and X_a. X_a is an increasing function of adhesion. Qualitatively, the influence of the surface layer on the thermal factor ψ_θ is similar to its influence on X_a. For instance, when other conditions are kept the same, the thickening of the surface layer decreases the chance for the sheared-off peak to adhere to the opposite surface by diffusion during temperature flash.

Consider the extreme case of vacuum in which two perfectly clean surfaces can join together by metallic adhesion. The adhesion factor of the chance of metal transfer X_a is, therefore, equal to one and will always be greater than or equal to the thermal factor ψ_θ. It is obvious that one should use Eqs. (1.02_b) and (1.10_b) in this case. They will have the following form:

$$M_t = M_t^*\phi_m,$$

and

$$M_w = 0.$$

This is equivalent to the statement that, in rubbing two perfectly clean surfaces in vacuum, every sheared-off blob of metal is transferred to the opposite surface, and no loose-wear particle will be formed.

In the other extreme case that the surface layer prevents adhesion entirely, i.e., X_a is zero, Eqs. (1.02_a) and (1.10_a) should always be used:

$$M_t = M_t^*\phi_m\psi_\theta \quad \text{when} \quad \psi_\theta > X_a, \qquad (1.02_a)$$

and

$$M_w = M_t^*\phi_m(1-\psi_\theta) \quad \text{when} \quad \psi_\theta > X_a. \qquad (1.10_a)$$

The most commonly encountered practical problems of sliding one metallic surface over another in ordinary atmospheric air are closer to this case than to the first one mentioned. Because of the very rapid initial rate of thickening of the oxide layer formed on most metallic surfaces, the surface layers formed in air, though very thin, are usually thick enough to prevent adhesion

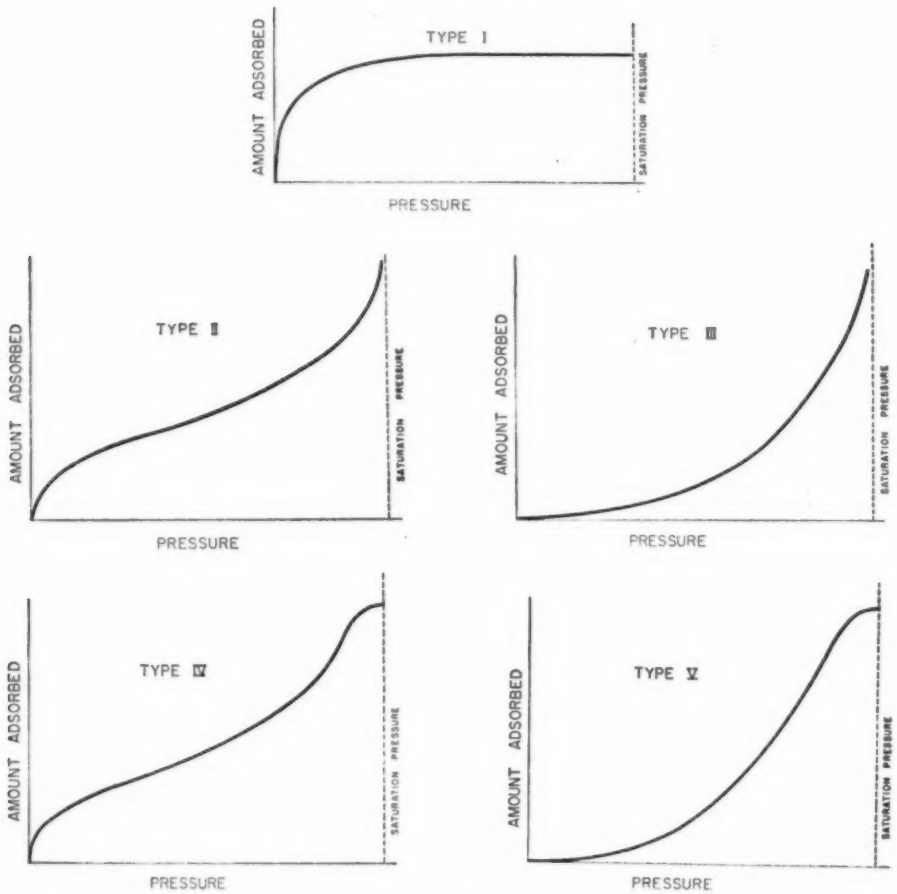


FIG. 2. The five isotherm types in physical adsorption. (Brunauer).

under the conditions of moderate load. Under heavy load, a certain amount of metal-to-metal contact could be established during the roughening of the interface if the surface film is thin and brittle. However, increase in load also raises the peak of the temperature flash which, in turn, pushes ψ_θ to a higher value. Only when the bulk temperature is extremely low and the sliding speed is very small, X_a may become greater than ψ_θ under heavy load. In other words, Eqs. (1.02_b) and (1.10_b) are used only under very special conditions.

In the formulation discussed under subsequent publications, Eqs. (1.02_a) and (1.10_a) will be used, and the case in which X_a is greater than ψ_θ will be treated as a special case.

PHYSICAL ADSORPTION

In physical adsorption, the gas molecules are probably adsorbed as quickly as they can reach the surface, because the adsorption reaches its equilibrium with such great rapidity that the rate of adsorption can hardly be measured. After the adsorption equilibrium is established, time has no more effect on the amount of adsorption. If the surrounding atmosphere causes only physical adsorption, metal transfer and wear are practically independent of the time of exposure to the atmosphere.

BOILING POINT

The data obtained by Hene[9] showed the parallelism between the boiling points and the amount of adsorption, other conditions being kept constant. The greater amount of adsorption for gases having higher boiling points will give a smaller value for both X_a and ψ_θ. Thus, the amount of wear will increase with boiling point of the gas, and the reverse will be true for the amount of metal transfer. The following list of most commonly encountered gases is given in the order of increasing boiling point, that is equivalent to the order of increasing wear or decreasing metal transfer:

Gas	Boiling point °C
H_e	−269
N_2	−196
O_2	−183
CO_2	− 78.5
H_2O	+100

It is to be remembered that this list applies only to the cases where the gases are merely physically adsorbed.

INERT GASES

Because of the inert behavior of the rare gases, they can only be physically adsorbed at low temperatures.

[9] W. Hene, Dissertation, University of Hamburg, 1927.

Among the rare gases, helium would be expected to have the simplest behavior. The attraction between gas molecules and the metal atoms on the solid surface is so feeble that physical adsorption becomes negligibly small at temperatures above about −80°C. Wear experiments in high vacuum are usually quite difficult. From the foregoing discussion, it can be seen that experiments carried out in a helium atmosphere at room temperature would be practically equivalent to experiments in high vacuum.

TEMPERATURE

Both temperature and gas pressure will influence metal transfer and wear. Since the equilibrium amount of physical adsorption always decreases with increasing temperature, the influence of temperature on X_a and

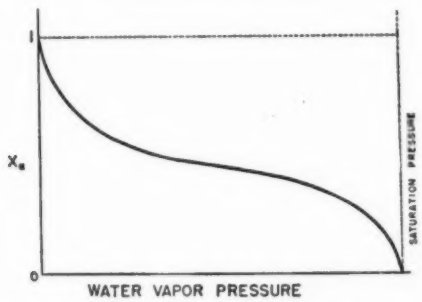


FIG. 3. Effect of water-vapor pressure on X_a.

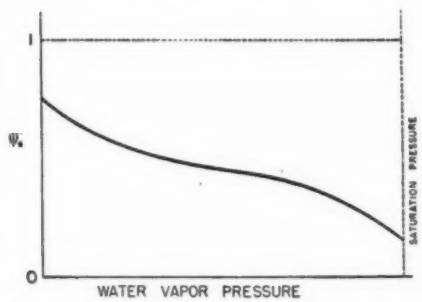


FIG. 4. Effect of water-vapor pressure on ψ_θ.

ψ_θ is to make them larger at high temperatures. The temperature influences metal transfer and wear not merely through its effect on the thickness of the physically adsorbed layer, for temperature also changes many other physical properties of metal. Therefore, it is very difficult to verify independently this effect of temperature on metal transfer and wear caused by change in physical adsorption.

The effect of gas pressure is opposite to that of temperature. High gas pressure will increase the adsorption at equilibrium, and hence, reduce X_a and ψ_θ.

ISOTHERM

The adsorption of water vapor by a metallic surface will serve as a good illustrative example because there is almost always some moisture in the ordinary atmosphere. Brunauer[10] showed five different types of adsorption isotherm (Fig. 2). Bowden and Throssell[11] have shown that the adsorption isotherm for water vapor on various metallic surfaces such as gold is of the Type II. The effect of water-vapor pressure on metal transfer and wear can be deduced qualitatively from the adsorption isotherm. Figures 3 and 4 show qualitatively the effect of pressure of an atmosphere of water vapor on X_a and ψ_θ, respectively. When the water-vapor pressure is zero, i.e. in vacuum, X_a is equal to one for perfectly clean surfaces, and ψ_θ is high but less than one. Both X_a and ψ_θ vary in the opposite way as the variation of adsorption with respect to the change in water-vapor pressure. Therefore, the shape of the X_a *vs* water-vapor pressure and ψ_θ *vs* water-vapor curves is more or less similar to the mirror image of the adsorption *vs* vapor-pressure curve. In the low-pressure region, X_a is expected to be larger than ψ_θ. In the high-vapor-pressure region, ψ_θ will in general become larger than X_a. Since the effect of a physically adsorbed layer on the mechanical interlocking is very small, the M_t *vs* water-vapor-pressure curve will have the same shape as the X_a *vs* water-vapor-pressure curve; or (as more likely) have the same shape as the X_a *vs* water-vapor-pressure curve in the low-pressure region and then follow the shape of the ψ_θ *vs* water-vapor-pressure curve in the high-pressure region where ψ_θ is usually greater than X_a (Fig. 5).

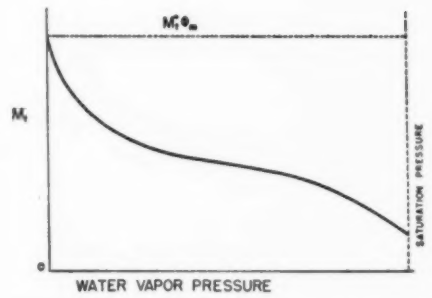


FIG. 5. Effect of water-vapor pressure on M_t.

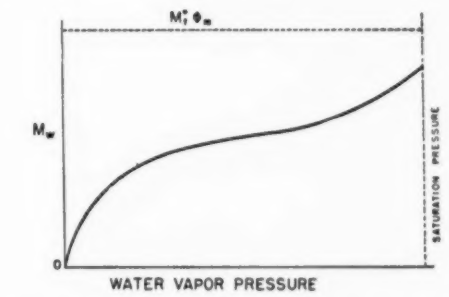


FIG. 6. Effect of water-vapor pressure on M_w.

[10] S. Brunauer, *The Adsorption of Gases and Vapors* (Princeton University Press, Princeton, 1945), Vol. I.

[11] F. P. Bowden and W. R. Throssell, Proc. Roy. Soc. (London) **A209**, 297–308 (1951).

The wear *vs* water vapor pressure curve‡ as shown in Fig. 6 is exactly the mirror image of metal transfer *vs* water vapor pressure curve and, hence, has a shape similar to the shape of Type II adsorption isotherm, except that M_w reaches a finite value rather than increasing to infinity when the water-vapor pressure approaches the saturation value.

By following the same way, a qualitative description of the effect of gas pressure on metal transfer and wear for any gas can be deduced from the isotherm when the shape of the isotherm for physical adsorption is known.

OXIDE-SURFACE LAYER

The oxide-surface film will be discussed as a representative case of chemical-compound layer because it is the most commonly encountered surface layer on metal surfaces. The thickness of oxide layer *vs* time of exposure curves obtained experimentally show a continuous increase in thickness accompanied by decreasing rate of thickening upon prolonged exposure to air or oxygen as illustrated in Fig. 1. Consequently, X_a will decrease with increasing time of exposure from one to practically zero (Fig. 7). The ψ_θ *vs* time of exposure curve shown in Fig. 8 starts from some value less than one at time zero and approaches zero asymptotically as the time increases towards infinity. The effect of the time of exposure to the oxidation on M_t and M_w as shown in Figs. 9 and 10 can be derived from the curves shown in Figs. 7 and 8.

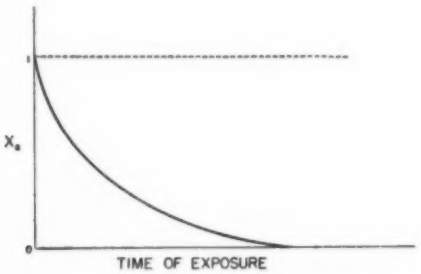


FIG. 7. Effect of time of exposure on X_a.

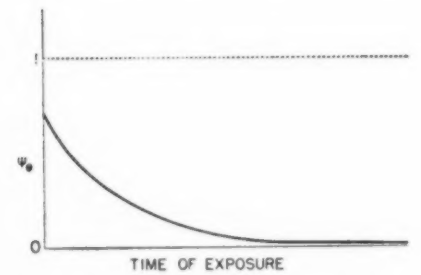


FIG. 8. Effect of time of exposure on ψ_θ.

‡ The shape of wear *vs* water vapor pressure curve agrees well with Feng and Chang's experimental result (to be published).

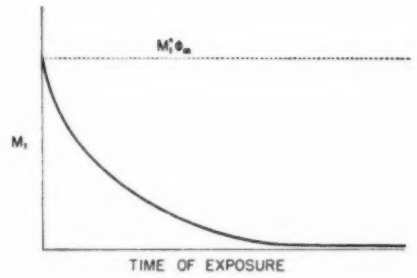


FIG. 9. Effect of time of exposure on M_t.

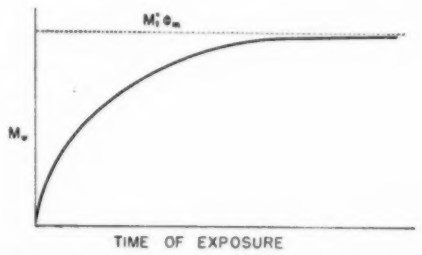


FIG. 10. Effect of time of exposure on M_w.

The exponential increase of the rate of oxidation with temperature is well known. Could it be made possible to separate the effect of temperature on X_a and ψ_θ caused by change in thickness of oxide layer from the effect of temperature on X_a and ψ_θ through its influence on physical properties of metal, both X_a and ψ_θ should be found to decrease with temperature. The thickness of oxide layer is also influenced by the oxygen pressure. The effect of the oxygen pressure on M_t and M_w can be deduced, if its effect on the thickness of the oxide layer is known.

JOURNAL OF APPLIED PHYSICS VOLUME 26, NUMBER 1 JANUARY, 1955

Cross-Wound Twin Helices for Traveling-Wave Tubes*

MARVIN CHODOROW AND E. L. CHU
Microwave Laboratory, Stanford University, Stanford, California
(Received November 23, 1953)

This paper describes a structure, namely, a cross-wound twin helix which overcomes the disadvantages of a conventional helix for high-voltage traveling-wave tubes. The disadvantages of a single helix suitable for high voltages are: (1) the impedance for electron interaction is reduced because of the energy content of the noninteracting space harmonics, and (2) the high impedance of some of the space harmonics can result in backward-wave oscillation. In a structure consisting of two helices wound in opposite directions, the symmetry of the fields results in most of the electric energy being stored in the fundamental component and most of the magnetic energy in the space harmonics. This results in a higher impedance for the fundamental component and a reduced impedance for the space harmonics. Typical numbers for dimensions suitable for 10-kv operation are an increase of a factor of 2 in the fundamental impedance and a reduction of a factor of about 20 in the -1 space harmonic for the twin helix as compared to the single helix.

1. INTRODUCTION

HELICES have been commonly used as propagating structures for traveling-wave tubes, and have been found to perform very satisfactorily for voltages of the order of 5000 volts or less. For this voltage range, the impedance of the helix ($E^2/\beta^2 P$, as defined by Pierce) is higher than that of most or all competing structures, and, of course, the band width is also very great. However, in power tubes with dimensions suitable for voltages somewhat above 5000 volts, one runs into difficulties with the ordinary helix. As the operating voltage (and velocity) is increased, one must increase the pitch of the helix, and in order to handle a high-power electron beam satisfactorily, one would also like to increase the radius a of the helix. This means that for high-voltage operation, say of the order of 10 000 volts, corresponding to a velocity of $0.2c$ or higher, one would like to design the helix with a value of $ka(=2\pi a/\lambda)$ of about 0.3 to 0.4 or larger.

For any helical structure of this kind, whether it has a large value of ka or not, the mode in which it is operated as a traveling-wave tube will have various space-harmonic components. These space-harmonic components travel at velocities different from the fundamental component, and if the beam velocity is selected to interact with the fundamental component, there will be no interaction with these higher space harmonics. However, these space harmonics do carry power, and for a helix with radii desirable for the voltages, such as mentioned above, it is found that a considerably larger fraction of the power is carried by the higher space harmonics, which represent useless field, than for a low-voltage tube with smaller helix. This additional power carried by the higher space harmonics represents a reduction in the impedance of the fundamental. These reduction factors have been calculated by Tien[1] with the result that the reduction in impedance can be quite appreciable even for ka as small as 0.2.

A further difficulty arising from the space harmonics which has been encountered more recently is the existence of so-called backward-wave oscillations discovered by Kompfner. One of the space harmonics, namely the $l=-1$ component, at a frequency corresponding approximately to $ka=0.5$, will have a negative velocity equal in magnitude to that of the fundamental component with which the electron beam is interacting. This means that the same mode propagating in the negative direction (i.e., group velocity opposite to beam velocity) will have a component which can synchronize with the beam and this leads to oscillation, the feedback being provided by the opposing directions of flow of beam and electromagnetic energy. Although the impedance of this space harmonic is comparable with that of the fundamental component, it has the property that its strong axial field is very close to the helix rather than on the axis, and therefore interaction with the beam can be avoided if the beam is small compared to the helix radius. Obviously, if the helix radius itself is small, this puts even more severe restrictions on the beam, and one cannot use a beam which comes close to filling the helix. If one does, then interaction occurs with the space harmonic and one gets the backward-wave oscillation. The presence of these oscillations then will make the design of a high-power amplifier extremely difficult.

Both the effects described, i.e., the reduction of the impedance of the fundamental and the existence of backward-wave oscillations make the use of a single-tape helix at voltages of the order of 10 kv or higher extremely difficult unless one is willing to go to extremely small helix radii and even smaller beam radii.

This paper discusses a structure, the twin helix, which overcomes both of these difficulties. The structure consists of two tape helices wound in opposite directions. Such a double helix has been investigated for other purposes, principally for applications where two modes of the structure are used simultaneously. These two modes (designated here as the symmetric and anti-

* The research reported in this paper was supported jointly by the U. S. Army Signal Corps, the U. S. Air Force, and the U. S. Navy (Office of Naval Research) under Office of Naval Research Contract No. N6onr-25123.

[1] P. K. Tien, "Helix Impedance of Traveling-Wave Tubes," Technical Report No. 50, Stanford University (June, 1952); also in Proc. Inst. Radio Engrs. 41, 1617 (1953).

symmetric modes) can be considered as arising from the combining of the single-helix modes of the two helices with different phases. In one case the two modes are superimposed in phase, and in the other, out of phase. The antisymmetric mode, which comes from superposition of the two single-helix modes 180° out of phase, results in a decreased axial field and a strong radial field between the helices. In the limit, as the helix diameters approach each other, the field on the axis will go to zero for this mode. The other mode (the symmetric mode, in which the two helix modes are superimposed in phase) results in an increased field on the axis, and very small radial field between them. Although many of the qualitative features of the two modes, which one gets in the twin helix, can be found by considering the simple superposition of the two single-helix fields, this does not give quantitative results since it ignores completely the interaction between the two helices. The phase velocities, for example, of each of the two modes differ quite markedly from the phase velocities of the mode for the single-tape helix. We discuss below some of the interesting properties which can be surmised from purely qualitative arguments, but the quantitative results given in this paper are for an actual twin-tape helix, taking into account the interaction of the two helices, and the numerical values should be quite accurate. As has been stated above, for the symmetric mode, in the limit, as the radii of the two helices become equal, the two helices can probably be allowed to touch with little perturbation of the mode since there is little radial field between the two helices. The antisymmetric mode, on the other hand, would presumably be strongly perturbed. In either case, some further investigation is necessary to find the extent of the perturbation. Since we are interested here principally in the symmetric mode for traveling-wave tube applications, we shall consider only this mode in detail. The methods used in the calculation, however, would be applicable to the antisymmetric mode.

Let us consider now what one might expect for the symmetric mode as a result of simple superposition of the fields of the two separate helices, using, for example, the expressions for the various space harmonics as given by Sensiper.[2] If one superimposes the fields in such phase that the axial electric fields of the fundamental component add, then many of the useful properties of the twin helix become apparent. First, one finds that if one does superimpose the fields of the two helices in this fashion, then H_z of the fundamental component is identically zero and all of the energy storage associated with H_z, which is useless as far as interaction with electron beams is concerned, is identically zero. If one examines the various terms in the stored energy for a single helix using the sheath model, it turns out that roughly half of the energy storage is associated with H_z and field components which are derived from it, i.e., with the TE portion of the field. This energy storage is mostly in the form of magnetic energy, whereas the energy associated with the TM portion of the field is mostly electric energy. For the sheath model, of course (and this will also be a good approximation to a low-velocity tape helix), one must necessarily have both the TM portion and the TE portion present in order to satisfy the boundary conditions. For a single-tape helix for which space-harmonic components are not negligible, it still turns out that in the fundamental component one has roughly equal amounts of TM and TE energy storage, or, what amounts to the same thing, electric and magnetic energy. For the twin-tape, from superposition, it becomes necessary for the space harmonics to carry principally magnetic energy, since there is no TE portion in the fundamental component. This is a symmetry property which should be preserved in a more exact analysis where one takes into account the interaction between the two helices. And indeed this turns out to be verified by the more detailed calculations, where it is found that the higher space harmonics do have most of their energy in TE portions of the field, which is mostly magnetic. This means that the space harmonics have relatively small axial components of electric fields, and therefore relatively small impedances for backward-wave oscillation. The details of the calculation and some numerical results are given in the body of the paper. It is found, for example, that for a twin helix suitable for use at 10 000 volts with $ka=0.4$ the impedance of the fundamental component is roughly two and one-half times as great as it is for a single-tape of the same diameter and designed for the same phase velocity. Contrariwise, the space harmonic impedances for the single-tape are much greater than they are for the double-tape. In particular, in the single-tape for the $l=-1$ space harmonic, the impedance is about 0.3 that of the fundamental, whereas for the double-tape it is only about 2 percent of the fundamental. One does have some disadvantage in the twin helix in that it is more dispersive and therefore would not be useful over as wide a band width for a fixed voltage. The same superposition approach to the antisymmetric mode indicates that there is essentially zero impedance for the fundamental component, most of the energy in the fundamental component being TE (magnetic), and conversely, there should be high impedances for the space harmonics where the energy would presumably be mostly electric.

2. THE MATHEMATICAL PROBLEM

The structure of twin helices is shown in Fig. 1. The helices are assumed to be infinitely long, of equal radii ($r=a$), wound with an infinitely thin and perfectly conducting tape.

The mathematical problem is simply to solve Maxwell's equations under appropriate boundary condi-

[2] S. Sensiper, "Electromagnetic Wave Propagation on Helical Conductors," Sc.D. thesis, Massachusetts Institute of Technology (1951); condensed version published as Mass. Inst. Technol. Research Lab. Electronics, Tech. Rept. No. 194, Massachusetts Institute of Technology.

tions. We shall only consider a source-free problem. The field intensities should be regular everywhere in space, and should satisfy the following boundary conditions:

$\mathbf{n}\times\mathbf{E}$ continuous everywhere on the cylindrical surface $r=a$, (2.1)

$\mathbf{n}\times\mathbf{E}=0$ on helices; (2.2)

$\mathbf{n}\times\nabla\times\mathbf{E}$ continuous on $r=a$ except on helices. (2.3)

The assumption that the helices are wound with an infinitely thin tape facilitates the solution of the problem. This would also facilitate a comparison between the results for twin helices and those for a single helix, because the latter case has been investigated in detail by Sensiper using the same assumption.†

3. SYMMETRY PROPERTIES OF THE STRUCTURE

The domain of the helix tapes is defined as

$$-\frac{\eta}{2}\leqq\left(\phi\mp\frac{2\pi z}{p}\right)\leqq\frac{\eta}{2}, \qquad (3.1)$$

$$\eta=\delta/a\,\tan\theta.$$

The minus sign is used for the right-handed helix, and the plus sign is for the left-handed helix.

The structure has several degrees of symmetry. The structure remains unchanged under any one or any combination of the following transformations:

$$(r,\phi,z)\rightarrow(r,\phi,-z), \qquad (3.2a)$$

$$(r,\phi,z)\rightarrow(r,\phi,z\pm p), \qquad (3.2b)$$

$$(r,\phi,z)\rightarrow(r,-\phi,z), \qquad (3.2c)$$

$$(r,\phi,z)\rightarrow\left(r,\phi\pm\pi,z+\frac{p}{2}\right). \qquad (3.2d)$$

Both the transformations (3.2a) and (3.2c) change the right-handed helix into a left-handed one and the left-handed into a right-handed, but do not change the whole structure. Equation (3.2a) represents reflection symmetry about the x-y plane; (3.2c) represents the reflection symmetry about the plane $\phi=0$ (z-x plane). Double helices of different radii do not have reflection symmetries. The invariant property of (3.2b) holds true for all periodic structures with periodicity p. The last invariant transformation, i.e., (3.2d), represents a unique property of double helices having the same pitch p regardless of radii being equal or different. This is in contrast with the case of a single helix which is invariant under the differential screw transformation

$$(r,\phi,z)\rightarrow\left(r,\phi\pm\frac{2\pi}{p}\delta z,z+\delta z\right),$$

† The distinction between insulated and noninsulated tapes will be considered in Sec. V. It does not affect the formulation of the problem at this point.

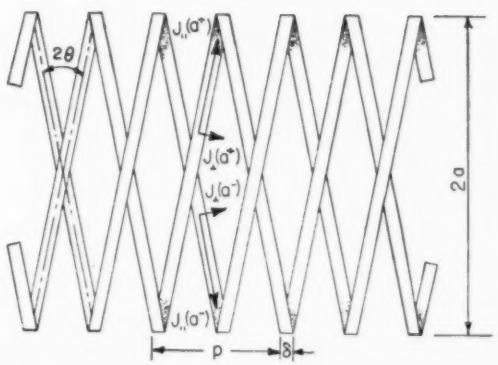


FIG. 1. Cross-wound twin helices. a is helix radius; p, pitch; θ, pitch angle; and δ, width of helix tape in the axial direction. $\cot\theta=2\pi a/p$, $\eta=2\pi\delta/p$. Helix tapes are assumed to be infinitely thin. The positive directions of the $\parallel$ and $\perp$ components of current on helix tapes are indicated.

where the choice of sign depends on whether the single helix is left or right-handed. Twin helices are invariant under this transformation only for $\delta z=p/2$.

When these reflection symmetries, (3.2a) and (3.2c), are considered together with those of Maxwell's equations, one finds that any solution of the problem must be either even or odd in ϕ and z except when there is degeneracy. In such case, although odd and even solutions exist they can be combined to give solutions which have neither symmetry. Thus in this problem the solutions which are odd or even in z represent standing waves. These obviously can be combined to give running waves. In Table I are listed the four different possible combinations of even and odd solutions which are compatible with Maxwell's equations and the symmetry properties. Since we are interested in the modes which have E_z on the axis, the solutions we shall look for are of types 1 and 2. These will be called the symmetric fields. Either by itself would give a standing wave, since the z dependence is either odd or even; to get a running wave we must combine solutions of the type 1 and 2. It is especially to be noted that for type 1 and 2, since H_z is odd in ϕ the fundamental Fourier component of H_z which has no dependence on ϕ must be identically zero. Corresponding statements could be made about the fundamental component of E_z in the antisymmetric mode (types 3 and 4).

4. METHOD OF SOLUTION

Although this is a vector problem, any field quantity can be expressed in terms of only two scalar functions, one specifying the TM field, the other the TE field in a

TABLE I. Symmetry of characteristic field vectors.

	Even in ϕ		Odd in ϕ	
Type	Even in z	Odd in z	Even in z	Odd in z
1	E_z, H_ϕ	E_r	H_r	H_z, E_ϕ
2	E_r	E_z, H_ϕ	H_z, E_ϕ	H_r
3	H_r	H_z, E_ϕ	E_z, H_ϕ	E_r
4	H_z, E_ϕ	H_r	E_r	E_z, H_ϕ

slow wave structure. For a convenient choice of these scalar functions, one may either take two components of field intensities E_z, H_z or two components of the vector potential or of the Hertz vector, or perhaps two components of the surface current density on one of the two helix tapes. The current on the other tape will be determined then by symmetry. As no simple known function will satisfy both the differential equations and the boundary conditions, we can construct a solution by the usual technique of superimposing an infinite number of elementary waves, each of which satisfies the differential equation and perhaps some of the symmetry conditions. The aggregate of waves will then satisfy Maxwell's equations, and by suitable choice of constants, one can also satisfy the boundary conditions and other symmetry conditions.

For twin helices, a suitable solution for, say E_z or H_z, would be a doubly infinite series of the form

$$\sum_{l=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} A_{l,m}R_{l,m}(r) \times\exp\left(j\left\{\left[\beta_{0,0}+\frac{2\pi}{p}(l+2m)\right]z-l\phi\right\}\right). \quad (4.1)$$

The functions $R_{l,m}(r)$ are then required to satisfy a particular differential equation. On a surface $r=$constant any infinite Fourier series of this sort can represent any function by a suitable choice of the coefficients $A_{l,m}$. Once the value of the function on this surface is known, the value everywhere is uniquely determined. Thus this series is a sufficient set of waves in terms of which the solutions of the problems can be constructed. For a single helix, one gets the same series, except that the index m is no longer required, and one gets a singly infinite series.

If one considers the general case of two different radii, of the helices a and b, one has different series of the form (4.1) for the regions $r<a$, $a<r<b$, and $r>b$. Thus one finds, for example for region $I(r<a)$

$$R_{l,m}{}^{I}(r)=I_l(\gamma_{l,m}r),$$

where $I_{l,m}$ is the so-called modified Bessel function and

$$\gamma_{l,m}{}^2=\beta_{l,m}{}^2-k^2,$$

$$\beta_{l,m}=\beta_{0,0}+\frac{2\pi}{p}(l+2m).$$

For region III, one has corresponding function $K_{l,m}(\gamma_{l,m}r)$ and for region II, combinations of the I's and K's. The series for any quantity in two neighboring regions should match each other on the common boundaries, except on the helices where E_r, H_ϕ, and H_z will have discontinuities depending on the current or charge density.

If we introduce surface current densities $J(a)$ and $J(b)$ and their components $J_\phi(a)$, $J_z(a)$, etc., one can get a set of obvious equations for the difference between various components of H on both sides of the tapes in terms of these current densities. If we expand these current densities in a Fourier series of the form

$$\mathbf{J}(a)=\sum_{l,m}\mathbf{J}(a)_{l,m}\exp[j(\beta_{l,m}z-l\phi)],$$
$$\mathbf{J}(b)=\sum_{l,m}\mathbf{J}(b)_{l,m}\exp[j(\beta_{l,m}z-l\phi)], \quad (4.2)$$

then all field quantities can be expressed in terms of these coefficients for the surface current densities. For the case $a=b$, one gets expressions for E_z and H_z as follows:

$$H_z=\sum_{l,m}(H_z)_{l,m}\exp[j(\beta_{l,m}z-l\phi)],$$
$$(H_z)_{l,m}=-(\gamma_{l,m}a)\cdot\begin{pmatrix}I_l'(\gamma_{l,m}a)K_l(\gamma_{l,m}r)\\ K_l'(\gamma_{l,m}a)I_l(\gamma_{l,m}r)\end{pmatrix}\cdot[J_\phi(a^-)_{l,m}+J_\phi(a^+)_{l,m}]; \quad (4.3)$$

$$E_z=\sum_{l,m}(E_z)_{l,m}\exp[j(\beta_{l,m}z-l\phi)],$$
$$(E_z)_{l,m}=\frac{1}{jka}(\gamma_{l,m}a)^2\cdot\begin{pmatrix}I_l(\gamma_{l,m}a)K_l(\gamma_{l,m}r)\\ K_l(\gamma_{l,m}a)I_l(\gamma_{l,m}r)\end{pmatrix}\cdot\left\{[J_z(a^-)_{l,m}+J_z(a^+)_{l,m}]-\frac{l\beta_{l,m}a}{(\gamma_{l,m}a)^2}\cdot[J_\phi(a^-)_{l,m}+J_\phi(a^+)_{l,m}]\right\}. \quad (4.4)$$

In these equations a^- and a^+ designate the right-handed and left-handed helices, respectively.

From E_z and H_z all other field quantities can be derived. It is to be noticed from the equations that tangential E will be continuous in crossing the tape, independent of the choice of the coefficients or the current amplitudes. The condition (2.3) is also satisfied provided that the coefficients in (4.2) are so chosen that the current amplitudes are zero off the helices. To satisfy Eq. (2.2), the field at $r=a$ given by (4.4) must be zero on the helices, and the corresponding value for E_ϕ must also be zero.

To meet these conditions we expand the current density as given in (4.2) in another series of functions $g_{\mu,\nu}$ which are orthogonal on the tape and zero off, e.g.,

$$\mathbf{J}(a)=\begin{cases}\sum_{\mu,\nu}\mathbf{J}(a)_{\mu,\nu}g_{\mu,\nu}(a,\phi,z) & \text{on helix}\\ 0 & \text{off helix.}\end{cases} \quad (4.2')$$

By multiplying this equation and the corresponding equation of (4.2) by $\exp[-j(\beta_{l,m}z-l\phi)]$ and integrating over a length p we can express the $\mathbf{J}(a)_{l,m}$ in terms of the $\mathbf{J}(a)_{\mu,\nu}$. The $\mathbf{J}(a)_{\mu,\nu}$ can be inserted into the expressions for E_z, E_ϕ, at $r=a$. Since these are required to be zero

on the tape, if we multiply, say, $E_z(a)$ by $g_{\mu\nu}(a)$ and integrate over the tape for a period p, the resulting algebraic equation must be identically zero for all μ, ν since E_z is zero on tape. The result is four doubly infinite sets of homogeneous algebraic equations [from $E_z(a)$, $E_z(b)$, $E_\phi(a)$, $E_\phi(b)$] to determine the coefficients $J_z(a)_{\mu,\nu}$, $J_\phi(a)_{\mu,\nu}$, $J_z(b)_{\mu,\nu}$, and $J_\phi(b)_{\mu,\nu}$. From these one can at least formally obtain a determinantal equation to determine ka as a function of $\beta_{00}a$ or vice versa. One then solves the simultaneous equations to get the four doubly infinite sets of coefficients.

The formal method described here differs in important detail from the *formal* procedure used by Sensiper[2] in getting corresponding algebraic equations for the single tape. In particular, his procedure did not ensure the vanishing of the current density *off* the tape. This formal error does *not* invalidate the approximation used by Sensiper in getting numerical results, and indeed some results of ours for the single tape, to be described later, agree with his.

The above briefly outlined formal method of solution, though of academic value, is not quite convenient for deriving approximate solutions to yield numerical results. We shall use instead a variational method for the actual calculation. This method will also lead to the same infinite sets of equations we have just described, but it has the great advantage that it permits the use of successive approximations which in the final limit would converge to the exact solution. The first few approximate solutions can be quite simple and, besides, the approximate solutions will in most cases approach the exact solution in a definite direction, either from above or from below.

If one starts from one of the standard forms of Lagrangian for an electromagnetic field, one can find the variational expression for the present problem in terms of field intensities which satisfy Maxwell's equations and symmetry and periodicity conditions but not the boundary conditions on the surface ($r=a$) of helix tapes. This may be written in the following form:

$$I=\int_0^p dz\int_0^{2\pi} ad\phi\mathbf{n}\cdot[\mathbf{E}^{(1)}(a)^*\times\mathbf{H}^{(1)}(a)-\mathbf{E}^{(3)}(a)^*\times\mathbf{H}^{(3)}(a)]$$
$$=\int_0^p dz\int_0^{2\pi} ad\phi\mathbf{n}\cdot\mathbf{E}(a)^*\times[\mathbf{H}^{(1)}(a)-\mathbf{H}^{(3)}(a)],$$

i.e.,

$$I=\text{constant}\cdot\sum_{l,m}\{E_\phi(a)^*_{l,m}\cdot[J_\phi(a^-)_{l,m}+J_\phi(a^+)_{l,m}]+E_z(a)^*_{l,m}\cdot[J_z(a^-)_{l,m}+J_z(a^+)_{l,m}]\}. \quad (4.5)$$

The physical interpretation of I is very simple; it is the complex power which might be generated or absorbed by the cylindrical surface $r=a$. The solution of the problem is then given by

$$\delta I[\mathbf{J}(a^-)_{l,m};\mathbf{J}(a^+)_{l,m}]=0. \quad (4.6)$$

Here one must notice that the small variations $\delta\mathbf{J}(a^-)_{l,m}$ and $\delta\mathbf{J}(a^+)_{l,m}$ are not entirely arbitrary; $\delta\mathbf{J}(a^-)$ and $\delta\mathbf{J}(a^+)$, like $\mathbf{J}(a^-)$ and $\mathbf{J}(a^+)$ themselves, must vanish off the helices. When this is taken care of, the four doubly infinite set of simultaneous equations obtained from (4.6) are exactly the same as those obtained by the first method.

5. APPROXIMATE SOLUTIONS FOR THE SYMMETRICAL MODE

In this section we shall show how to apply (4.6) to get approximate solutions of our problem. Before doing this we should like to derive some useful symmetry properties of the current amplitudes.

As mentioned in Section 3, we shall only be concerned with the symmetrical solution. In this case E_r, E_z, and H_ϕ are even in ϕ, while H_r, H_z, and E_ϕ are odd in ϕ. From the relations between the components of H and J it follows that $J_z(a)$ should be even and $J_\phi(a)$ odd in ϕ; in other words,

$$J_z(a^+,\phi,z)=J_z(a^-,\phi,z),$$
$$J_\phi(a^+,\phi,z)=-J_\phi(a^-,-\phi,z). \quad (5.1)$$

Let $J_\parallel(a^-)$ and $J_\perp(a^-)$ be the components of the surface current density on the right-handed helix in directions respectively parallel and perpendicular to the edge of the helix tape; let $J_\parallel(a^+)$ and $J_\perp(a^+)$ be the corresponding currents on the left-handed helix. The two directions of each pair of currents ($J_\parallel(a^-)$, $J_\parallel(a^+)$ and $J_\parallel(a^-)$, $J_\parallel(a^+)$) are symmetrically oriented with respect to the z axis; as shown in Fig. 1. J_z and J_ϕ may be expressed in terms of $J_\parallel$ and $J_\perp$, and the angle θ in an obvious way

$$\begin{matrix}J_z(a^-)\\J_\phi(a^-)\end{matrix}=J_\parallel(a^-)\begin{matrix}\sin\theta\\\cos\theta\end{matrix}\pm J_\perp(a^-)\begin{matrix}\cos\theta\\\sin\theta,\end{matrix} \quad (5.2a)$$

$$\begin{matrix}J_z(a^+)\\-J_\phi(a^+)\end{matrix}=J_\parallel(a^+)\begin{matrix}\sin\theta\\\cos\theta\end{matrix}\pm J_\perp(a^+)\begin{matrix}\cos\theta\\\sin\theta,\end{matrix} \quad (5.2b)$$

and the condition (5.1) becomes

$$J_\parallel(a^+,\phi,z)=J_\parallel(a^-,-\phi,z),$$
$$J_\perp(a^+,\phi,z)=J_\perp(a^-,-\phi,z). \quad (5.1')$$

The Fourier coefficients of $J_\parallel$ and $J_\perp$ may be similarly defined:

$$J_{\parallel\atop\perp}(a^-,\phi,z)=\sum_{l,m}J_{\parallel\atop\perp}(a^-)_{l,m}\exp[j(\beta_{l,m}z-l\phi)], \quad (5.3a)$$

with similar expressions for $J_\parallel(a^+,\phi,z)$. Since

$$\beta_{l,m}=\beta_{-l,m+l}, \quad (5.4)$$

we obtain from (5.1′) and (5.3) the following symmetrical relation:

$$J_{\parallel\atop\perp}(a^+)_{l,m}=J_{\parallel\atop\perp}(a^-)_{-l,m+l}. \quad (5.5)$$

By using the above relations, we find from (4.3)

$$(H_z)_{l,m}=0 \quad \text{unless} \quad l\neq 0. \tag{5.6}$$

Thus the symmetrical solution has no H_z component with $l=0$, and this is rigorously true with no approximation involved. In other words, the fundamental component of the wave field is a pure TM field; the TE parts of the fundamental component arising from the two helices cancel each other.

The analysis can be very much simplified by assuming that

$$J_\perp \cot\theta \ll J_\parallel \quad \text{and} \quad J_\perp \tan\theta \ll J_\parallel.$$

This assumption should be good at least for narrow tapes, and using it we can drop all the $J_\perp$ terms in (5.2).

By writing out the variational expression $I[J(a^-)_{l,m}; J(a^+)_{l,m}]$ and using the relations (5.2) with the $J_\perp$ terms omitted, and simplifying, we finally obtain I as an infinite bilinear sum of the $J_\parallel(a^-)_{l,m}$ with coefficients involving $K_l(\gamma_{l,m}a)$, $I_l(\gamma_{l,m}a)$, their derivatives, $\cos\theta$, $\sin\theta$, etc. The variation of I is to be carried out under the condition that

$$\sum_{l,m} J_\parallel(a^-)_{l,m} \exp[j(\beta_{l,m}z-l\phi)]=0 \text{ off helix.} \tag{5.7}$$

To insure this, one needs to transform I by expressing $J_\parallel(a^-)_{l,m}$ in terms of some other set of coefficients, which *can* be varied independently. We can introduce functions $g_{\mu,\nu}(a^-,\phi,z)$ as in (4.2′), which form a complete orthogonal set defined in the domain of the tape (a^-), and then write any admissible helix current

$$J_\parallel(a^-)=\sum_{l,m} J_\parallel(a^-)_{l,m} \exp[j(\beta_{l,m}z-l\phi)]$$

as

$$J_\parallel(a^-)=\begin{cases}\sum_{\mu,\nu} J_\parallel(a^-)_{\mu,\nu}\cdot g(a^-)_{\mu,\nu} & \text{on helix}\\ 0 & \text{off helix.}\end{cases} \tag{5.8}$$

Thus $J_\parallel(a^-)_{l,m}$ can be expressed in terms of $J_\parallel(a^-)_{\mu,\nu}$. Varying the coefficients $J_\parallel(a^-)_{\mu,\nu}$ varies the current on the helix but not the value off helix, so $J_\parallel(a^-)_{\mu,\nu}$ can be varied independently.

For an approximate solution we may assume $J_\parallel(a^-)$ to consist of a finite number of terms of certain convenient functions, each term having an unknown coefficient. For a given sequence, increasing the number of terms increases the degree of accuracy. The sequence functions do not have to be orthogonal functions in the domain of the helix tape; in fact, we can just take the sequence $\exp[j\beta_{\mu,\nu}z\mp l\phi)]$, which we find is quite convenient for calculation. Thus we assume

$$J_\parallel(a^-)=\begin{cases}\sum_{\mu=l'}^{l''}\sum_{\nu=m'}^{m''} A_{\mu,\nu}^- \exp[j(\beta_{\mu,\nu}z-\mu\phi)] \\ \quad +A_{\mu,\nu}^+ \exp[j(\beta_{\mu,\nu}z+\mu\phi)] & \text{on helix}\\ 0 & \text{off helix,}\end{cases} \tag{5.9}$$

and we can evaluate the $J_\parallel(a^-)_{l,m}$ by integration in z over a length p. In this connection, one may note that only one of the two series in (5.9) with $\nu=0$ is admissible for the case of a single helix, $A_{\mu,0}^-$ series for a right-handed and $A_{\mu,0}^+$ series for a left-handed helix. In the case of twin helices both series are admissible.‡

For most practical purposes, only a few terms are needed in the series for the assumed current. This is known from the experience with similar problems, especially the problem of a single helix. Sensiper assumed a one-term current $e^{j\beta_0 z}$ and obtained reliable results for the single helix. For a one-term approximation, there is no variational problem; the determinantal equational is simply

$$I=0. \tag{5.10}$$

We have considered the following single-term and two three-term approximations:

on helix

$$J_\parallel(a^-)=A\exp(j\beta_{0,0}z), \tag{5.11a}$$

$$J_\parallel(a^-)=A\exp(j\beta_{0,0}z)\cdot\left\{1+f\exp\left[j\left(\frac{2\pi z}{p}-\phi\right)\right]+h\exp\left[-j\left(\frac{2\pi z}{p}-\phi\right)\right]\right\}, \tag{5.11b}$$

$$J_\parallel(a^-)=A\exp(j\beta_{0,0}z)\cdot\left\{1+f'\exp\left[j\left(\frac{2\pi z}{p}+\phi\right)\right]+h'\exp\left[-j\left(\frac{2\pi z}{p}+\phi\right)\right]\right\}, \tag{5.11c}$$

and off helix

$$J_\parallel(a^-)=0.$$

In passing, we may note that if we only use $A_{\mu,\nu}^-$ series in Eq. (5.9) as in (5.11b)

$$J_\parallel(a^-)_{l,m}=0 \quad \text{unless } m=0.$$

‡ It is at this point in the analysis that it is necessary to make a distinction between insulated tapes and touching tapes. For insulated tapes, one must have continuity of current on each tape, and in particular, one cannot have current flowing off the edge of the tape. For touching tapes, however, one can have a current distribution corresponding to current flowing off the tape, in the region where the two tapes overlap, since a corresponding current can be assumed on the other tape to provide continuity. If the overlap region is very small, corresponding to very narrow tapes, the possibility or impossibility of currents flowing from one tape to the other will make very little difference, and an assumed current which has no flow from one tape to the other should be quite good even if the tapes are touching. For wide tapes, however, with a correspondingly large region of overlap, current paths can be appreciably shorter by flow from one tape to the other; and, therefore, the behavior of the structure will be different, depending on whether such current flow is possible or not. Any assumed current distribution which does not have flow from one tape to the other would be good only for insulated tapes. Since all the current distributions used in this paper are of a kind which have no flow between tapes, it is to be expected that the results would not be too accurate for wide tapes in contact. Experiments performed at Stanford University by Mr. John E. Nevins, Jr., on wide tapes in contact or separated by 0.001-in. polystyrene show that in the latter case one got much better agreement with the theory than in the former.

So, with this assumption, there is only a singly infinite set of coefficients $J_{\mathrm{II}}(a^-)_{l,0}$, and if we wish, we may avoid the use of double subscripts. However, we shall keep the subscript $m=0$ for such cases in order to avoid confusion with the case of a single helix. It is only when we admit terms in the $A_{\mu,\nu}{}^{+}$ series that the use of double subscripts is essential.

So far most of the numerical work has been done on the one-term approximation; only a few cases have been checked by three-term approximations. These cases indicate that the one-term approximation does give reliable quantitative results for twin helices as for the single helix. In the following sections we shall discuss various properties of twin helices based exclusively on the one-term approximation. For that case we get

$$J_{\mathrm{II}}(a^-)_{l,m}=A\frac{\eta}{2\pi}\mu_l\delta(m), \tag{5.12}$$

$$\mu_l=\sin(l\eta/2)/(l\eta/2),$$

$$\delta(m)=\begin{cases}1 & m=0,\\ 0 & m\neq 0.\end{cases} \tag{5.13}$$

$\beta_{0,0}a$ is obtained from the determinantal equation

$$I=2U_{0,0}+\sum_{l=1}^{\infty}\mu_l^2(Y_{l,0}+Y_{-l,0})=0, \tag{5.14}$$

where

$$U_{0,m}=(\gamma_{0,m}a)^2\cdot K_0(\gamma_{0,m}a)I_0(\gamma_{0,m}a)\sin^2\theta, \tag{5.15}$$

$$Y_{l,m}=(\gamma_{l,m}a)^2\cdot K_lI_l\sin^2\theta$$
$$+\left[\left(\frac{l\beta_{l,m}a}{\gamma_{l,m}a}\right)^2K_lI_l+(ka)^2K'_lI_l'\right]\cos^2\theta$$
$$-(l\beta_{l,m}a)K_lI_l\sin 2\theta. \tag{5.16}$$

6. PHASE VELOCITY AND GROUP VELOCITY

With a given value of θ and η, the determinantal Eq. (5.14) can be solved numerically to obtain ka as a function of $\beta_{0,0}a$, or vice versa. The numerical calculations for the series in (5.14) were made by subtracting the terms of a known series, and a similar approach can be used for the corresponding determinant for the single helix. Calculations have been carried through for several cases for both the single and twin helices.

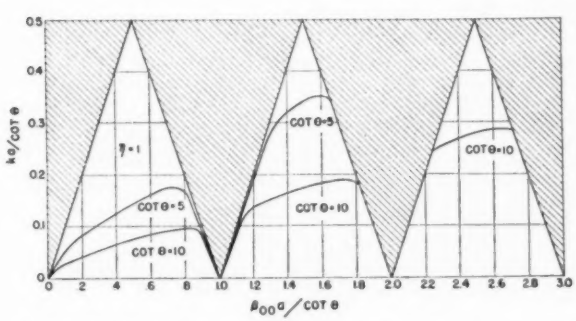


FIG. 2. $ka/\cot\theta$ *vs* $\beta_{0,0}a/\cot\theta$ for twin-helix structures. There are two branches for $\cot\theta=5$, $\eta=1$, and five branches for $\cot\theta=10$, $\eta=1$ (last two branches not shown). No solution to the source-free problem exists in the shaded region, in which $(\beta_{0,0}a+l\cot\theta)^2-(ka)^2>0$ is not satisfied for at least one value of l, l being integers, $-\infty\leq l\leq\infty$.

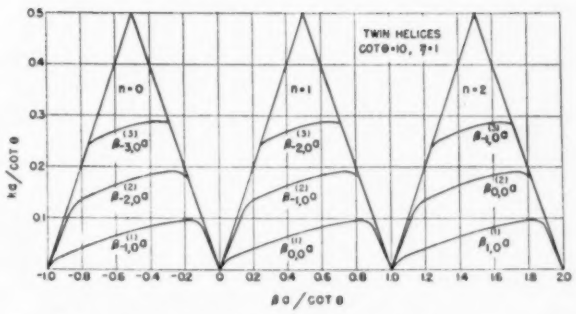


FIG. 3. $ka/\cot\theta$ *vs* $\beta a/\cot\theta$ for twin-helix $\cot\theta=10$, $\eta=1$. n denotes the number of a triangular pass region; n increases in the same direction as βa. In each pass region there are different branches (only the first three are shown) of solution pertaining to different components or space harmonics. $\beta_{l,0}{}^{(i)}a$ indicates the ith branch of the lth harmonic. i increases in the same direction as ka. $l+i=n$ in any one pass region.

In Fig. 2, $ka/\cot\theta$ is plotted against $\beta_{0,0}a/\cot\theta$ for two cases of twin helices, one having $\cot\theta=5$, $\eta=1$ and the other $\cot\theta=10$, $\eta=1$. The shaded area is the so-called "forbidden region" in which not all $\gamma_{l,0}$'s are real and so no solution to the source-free problem would exist.[2] The forbidden region for twin-helix structures is the same as that for the single helix. The solution for the case $\cot\theta=5$ has two different branches, whereas the solution for $\cot\theta=10$ has five (only three are shown in the figure). As the magnitude of $\cot\theta$ is kept increasing, there will be more and more different branches. These curves have the same general shape as the corresponding ones for the single helix. This is not surprising because the determinantal equations for the two structures are quite similar.

As indicated in Fig. 3, the order of a branch is designated by a superscript; $\beta_{0,0}{}^{(i)}a$ designates the ith branch of the solution for $\beta_{0,0}a$, $i=1, 2, 3, \cdots$, arranged in increasing order of frequency. To each branch $\beta_{0,0}{}^{(i)}a$ of $\beta_{0,0}a$ there is a corresponding one $\beta_{l,0}{}^{(i)}a$ of $\beta_{l,0}a$; $\beta_{l,0}{}^{(i)}a=\beta_{0,0}{}^{(i)}a+l\cot\theta$. Thus, the curves for $\beta_{-1,0}{}^{(2)}a$, $\beta_{-2,0}{}^{(3)}a$, $\cdots$ lie in the same triangular region as $\beta_{0,0}{}^{(1)}a$, and those for $\beta_{-1,0}{}^{(3)}$, $\beta_{-2,0}{}^{(4)}\cdots$ lie in the same triangular region as $\beta_{0,0}{}^{(2)}a$, etc. Three lower branches are shown in each triangular region in Fig. 3 for the case $\cot\theta=10$, $\eta=1$.

Among the various branches of the solution, the first branch is of most practical interest. For brevity we shall subsequently omit the superscript (1) for this branch. The phase velocity of the fundamental component $ka/\beta_{0,0}a$ for two other cases, $\cot\theta=2.5$, $\eta=1$ and $\cot\theta=20$, $\eta=1$, have also been calculated; they are plotted *versus* ka in Fig. 4 together with the two previous cases and two single helix cases ($\cot\theta_1=2.5$ and 4.785, $\eta=1$).

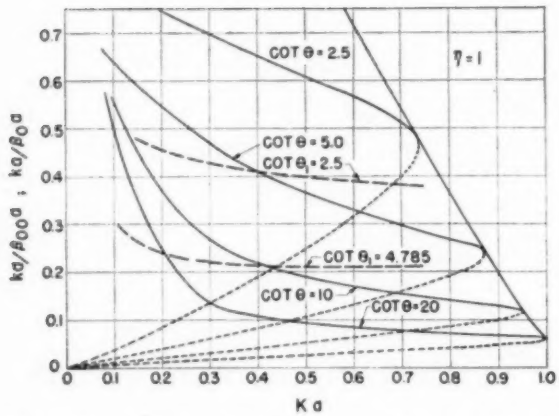


FIG. 4. Phase velocity ($ka/\beta_{0,0}a$) curves for twin-helix structures ($\eta=1$; $\cot\theta=2.5$, 5, 10, and 20). Also shown are sections of phase velocity ($ka/\beta_0 a$) curves for two single-helix cases ($\eta=1$; $\cot\theta_1=2.5$ and 4.785). They are drawn as dashed lines for reason of distinctness. The dotted part of twin-helix curves represent waves with negative group velocity.

From the curves shown in Fig. 3 one obtains the group velocity of a given component, which is defined as $d(ka)/d(\beta a)$. In Fig. 5, the group velocity v_g/c of the first branch is plotted against ka for the same four cases shown in the last figure.

In this connection we may mention that, if $\cot\theta\gg1$, simple accurate formulas can be derived from the determinantal equation for calculating the phase velocity and the group velocity over the whole range of variables. In particular, the asymptotic formulas for $\gamma_{0,0}a\ll1$ or $\gg1$ or anywhere near the forbidden region are very simple regardless of the value of $\cot\theta$. The same can be said of the determinantal equation for the single helix.

We note in Fig. 4 that with the same pitch angle and η, the phase velocity at given ka is greater for twin helices, and in Fig. 5 that the group velocity in the case of twin helices changes more rapidly as ka changes. The single helix has less dispersion and wider band width. On the other hand, we note from Figs. 4 and 5 that on the basis of the same phase velocity and the same ka the group velocity is smaller for twin helices than for the single helix having the same η. This reduction in group velocity tends to reduce the power flow, thus increasing the impedance parameter which is inversely proportional to the power. These changes, however, are only of moderate magnitude and the wide-band characteristic of the helix will not be changed greatly.

As regards the effect of changing η on the phase velocity, calculation shows that v_p increases with η and the effect is much smaller for the single helix than for twin helices.

7. FIELD INTENSITIES, ENERGY DENSITY, ENERGY FLOW, AND IMPEDANCE PARAMETERS

To facilitate the comparison with single helix, we transform the expressions of E_z and H_z, (4.3) and (4.4) into the following forms by using relations (5.4) and the symmetry properties (5.5):

$$E_z=\frac{1}{jka}\sum_{l,m}[(\gamma_{l,m}a)^2J_z(a^-)_{l,m}-(l\beta_{l,m}a)J_\phi(a^-)_{l,m}]\cdot\frac{I_l(\gamma_{l,m}a)K_l(\gamma_{l,m}r)}{K_l(\gamma_{l,m}a)I_l(\gamma_{l,m}r)}\cdot\exp(j\beta_{l,m}z)\cdot[\exp(jl\phi)+\exp(-jl\phi)], \quad (7.1)$$

and

$$H_z=\sum_{l,m}J_\phi(a^-)_{l,m}\cdot(\gamma_{l,m}a)\cdot\frac{I_l'(\gamma_{l,m}a)K_l(\gamma_{l,m}r)}{K_l'(\gamma_{l,m}a)I_l(\gamma_{l,m}r)}\cdot\exp(j\beta_{l,m}z)\cdot[\exp(jl\phi)-\exp(-jl\phi)]. \quad (7.2)$$

The corresponding expressions for the case of a single helix may be obtained from (7.1) and (7.2) by deleting the term $\exp(jl\phi)$ and replacing $\beta_{l,m}$, $\gamma_{l,m}$ by β_l, γ_l.

From E_z and H_z all other field components can easily be derived from Maxwell's equations, and from the field components a straightforward integration will give the energy density. Let us denote

$$W\equiv\sum_{l,m}W_{l,m}=W_{l=0}+W_{l\neq0}. \quad (7.3)$$

If we assume $J_{II}(a^-)$ to have only terms in the $A_{\mu,\nu}^-$ group, all the nonvanishing coefficients $J_{II}(a^-)_{l,m}$ have $m=0$. Then,

$$W_{l=0}=W_{0,0}$$

and

$$W_{l\neq0}=2\sum_{l=1}^{\infty}(W_{l,0}+W_{-l,0}). \quad (7.4)$$

For the corresponding case of a single helix we have

$$W_I=\sum_l W_l=W_0+\sum_{l=1}^{\infty}(W_l+W_{-l}). \quad (7.5)$$

In (7.5), W_0, the energy density of the fundamental component, will have a part coming from the TE field

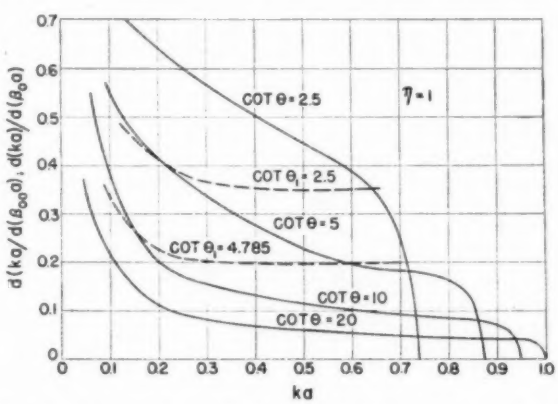


FIG. 5. Group velocity $[d(ka)/d(\beta_{0,0}a)]$ curves for twin-helix structures ($\eta=1$; $\cot\theta=2.5$, 5, 10, and 20). Also shown are sections of group velocity $[d(ka)/d(\beta_0 a)]$ curves for two single-helix cases ($\eta=1$; $\cot\theta_1=2.5$ and 4.785).

and another from the TM field. It is relevant to our later discussion to compare these two terms. We can show

$$\frac{(W_0)_{TE}}{(W_0)_{TM}}=\left(\frac{ka}{\gamma_0 a}\right)^2\cdot\cot^2\theta_1 \cdot\frac{[(\beta_0 a/ka)^2+1]K_1I_1+(\gamma_0 a)(K_0I_0'+K_0'I_0)}{[(\beta_0 a/ka)^2+1]K_0I_0+(\gamma_0 a)(K_0I_0'+K_0'I_0)} \cong K_1I_1/K_0I_0. \tag{7.6}$$

Thus the two parts of energy in the fundamental component, $(W_0)_{TM}$ and $(W_0)_{TE}$, are of the same order of magnitude except for small values of $\gamma_0 a$, i.e., $\gamma_0 a \ll 1$. We wish to compare this with the double helix.

On the basis of the same axial electric field, i.e., $(E_z)_{0,0}=(E_z)_0$, with same ka and same phase velocity, i.e., $\beta_{0,0}a=\beta_0 a$, and $\gamma_{0,0}a=\gamma_0 a$, we should have

$$2J_z(a^-)_{0,0}=J_z(a)_0$$

i.e.,

$$2J_{II}(a^-)_{0,0}\sin\theta=J_{II}(a)_0\sin\theta_1,$$

and this will give the same energy density for the TM field of the fundamental component for both structures. However, for the double helix since $(H_z)_{0,m}=0$ there is no TE part; therefore,

$$(W_{0,0})_{TM}\equiv W_{0,0}=(W_0)_{TM}$$

and

$$\frac{W_{0,0}}{W_0}=\frac{1}{1+[(W_0)_{TE}/(W_0)_{TM}]}\cong\frac{K_0I_0}{K_0I_0+K_1I_1}; \tag{7.7}$$

i.e., $W_{0,0}$ is about half of W_0.

It is also illuminating to compare the relative proportions of electric and magnetic energy stored in the fundamental components of both the single and twin helices. As defined, $W_{0,0}$ contains both the electric energy $(W_{0,0})_E$ and the magnetic energy $(W_{0,0})_H$; $W_{0,0}=(W_{0,0})_E+(W_{0,0})_H$.§

The ratio of the two parts of energy can be calculated

$$\frac{(W_{0,0})_H}{(W_{0,0})_E}=(2-\rho)\Big/\left[2\left(\frac{\gamma_{0,0}a}{ka}\right)^2+(2-\rho)\right], \tag{7.8}$$

where ρ is defined as

$$\rho\equiv-(\gamma_{0,0}a)\cdot(K_0I_0'+K_0'I_0)/K_0I_0,$$

$\rho\cong 1$ for $0.8\leqq\gamma_{0,0}a\leqq\infty$; in fact, $0\leqq\rho\leqq 1.06$ over the entire range $0\leqq\gamma_{0,0}a\leqq\infty$. Since $(\gamma_{0,0}a/ka)^2\gg 1-(\rho/2)$, we find that $(W_{0,0})_H$ is much smaller than $(W_{0,0})_E$; i.e., $W_{0,0}$, the energy of a pure TM field, is mainly electric.

§ Even though $(H_z)_{0m}\equiv 0$, there will be magnetic energy stored in the fundamental component, in the form of $(H_\phi)_{0m}$ which is a part of the TM field.

For the single helix it turns out

$$\frac{(W_0)_H}{(W_0)_E}\cong\frac{K_1I_1}{K_0I_0}. \tag{7.9}$$

$(W_0)_H$ is of the same order of magnitude as $(W_0)_E$ except for small values of $\gamma_0 a$. In the latter case the electric energy will also be predominant. The similarity between (7.6) and (7.9) is due to the fact that for a helix the TE part of W_0 is mainly magnetic and the TM part mainly electric.

One also obtains easily [with same normalization $(E_z)_{0,0}=(E_z)_0$]

$$\frac{(W_{0,0})_E}{(W_0)_E}\cong\left[1+\frac{(K_1I_1/K_0I_0)-(\rho/2)}{(\beta_0 a/ka)^2-(\rho/2)}\right]^{-1}<1, \tag{7.10a}$$

$$\frac{(W_{0,0})_H}{(W_0)_H}<\frac{1}{2}\cdot\frac{K_0I_0(1-\rho)}{K_1I_1\cot^2\theta_1}. \tag{7.10b}$$

The latter ratio is a very small quantity, because $\cot^2\theta_1\gg 1$.

It is a property of periodic structures that the total electric energy stored in the field in one periodic length must be equal to the total magnetic energy of the field in the same length. In fact, one can verify that the determinantal Eq. (5.14) is nothing but a statement that the above property must be true. As a result of this and (7.8), the space harmonics in the twin-helix structure must have greater magnetic energy than electric energy, whereas from (7.9) the energy in the space harmonics of the single-helix structure should divide approximately equally between electric and magnetic. It can be shown that, if the two structures also have the same η, the electric energy stored in any space harmonic in a twin-helix is less than half the electric energy stored in the corresponding harmonic in single-helix. The fact that the space harmonics in the twin-helix structure contain little electric energy implies that there is little axial electric field in these components, so the corresponding impedances will be low. This is borne out by numerical results to be described.

To compare the difference of energy densities of space harmonics in the two structures we shall first state an approximate relation between the pitch angles θ and θ_1. From the determinantal Eq. (5.14) we can obtain, by neglecting a remainder term which is small in comparison with other terms in the usual range of interest,

$$\left(\frac{ka\cot\theta}{\gamma_{0,0}a}\right)^2\cong\frac{2K_0I_0+f(\eta)\sin\theta}{f(\eta)\sin\theta}, \tag{7.11}$$

where

$$f(\eta)\equiv\frac{2}{\eta^2}\cdot[C_3(0)-C_3(\eta)]. \tag{7.12}$$

Where

$$C_3(\eta)=\sum_{l=1}^{\infty}\frac{\cos l\eta}{l^3}=1.202+(1/2)\eta^2\log\eta-(3/4)\eta^2,$$

$f(\eta)$ approaches infinity logarithmically as $\eta \to 0$; $f(\eta)$ varies slowly and is of the order of magnitude unity for practical values of η. Similarly, we obtain

$$\left(\frac{ka\cot\theta_1}{\gamma_0 a}\right)^2 \cong \frac{K_0 I_0 + f(\eta)\sin\theta_1}{K_1 I_1 + f(\eta)\sin\theta_1}. \tag{7.13}$$

From (7.11), (7.13), and the relations $\gamma_0 a = \gamma_{0,0} a$, $\beta_0 a = \beta_{0,0} a$, a quadratic equation for $x \equiv \csc\theta/\csc\theta_1$ can be obtained.

With $K_0 I_0 \cong 1/2\gamma_{0,0} a$ for relatively large values of $\gamma_{0,0} a$, we get approximately

$$x^2 \cong \frac{1}{f(\eta)\cdot(\gamma_{0,0} a/\cot\theta)} + 1. \tag{7.14}$$

From the various field expressions we can find the following relations among the various energy densities,

$$\frac{(W_{l,0})_H}{(W_{l,0})_E} \cong \frac{\csc^2\theta}{\csc^2\theta_1} = x^2, \tag{7.15a}$$

$$\frac{(W_l)_H}{(W_l)_E} \cong 1, \tag{7.15b}$$

$$\frac{2(W_{l,0})_E}{(W_l)_E} \cong \frac{1}{2}\cdot\frac{K_l(\gamma_{l,0} a) I(\gamma_{l,0} a)}{K_l(\gamma_l a) I_l(\gamma_l a)} \cong \frac{1}{2}\frac{\csc\theta_1}{\csc\theta} = \frac{1}{2x}, \tag{7.16a}$$

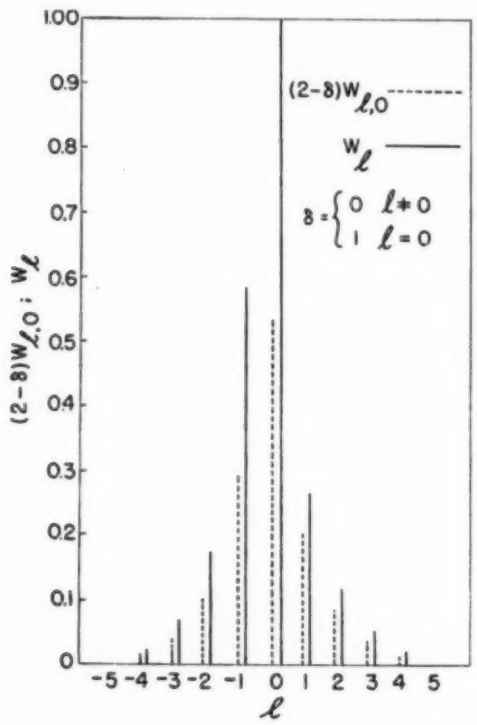


FIG. 6. Comparison of energy densities for the first few components of a single-helix ($\eta=1$, $\cot\theta_1=4.785$) and a twin-helix ($\eta=1$, $\cot\theta=10$). $ka=0.4$, $\beta_0 a=\beta_{0,0} a=1.888$. Arbitrary scale, normalized to same E_z on axis.

$$\frac{2(W_{l,0})_H}{(W_l)_H} \cong \frac{1}{2}\cdot\frac{K_l(\gamma_{l,0} a) I_l(\gamma_{l,0} a)}{K_l(\gamma_l a) I_l(\gamma_l a)}\cdot\frac{\csc^2\theta}{\csc^2\theta_1}$$

$$\cong \frac{1}{2}\cdot\frac{\csc\theta}{\csc\theta_1} = \frac{x}{2}. \tag{7.16b}$$

Since $\csc\theta/\csc\theta_1 > 1$, we find from (7.16a) $2(W_{l,0})_E/(W_l)_E < \frac{1}{2}$. This bears out the statement that the electric energy density of any space harmonic in a twin-helix is less than half of the corresponding energy density in the single helix. For the total stored energy by summing up the infinite series in W and W_I we find

$$\frac{W}{W_I} \cong \frac{1}{2}\cdot\frac{2K_0 I_0 + [(\csc\theta/\csc\theta_1) + (\csc\theta_1/\csc\theta)] f(\eta)\sin\theta_1}{K_0 I_0 + K_1 I_1 + 2f(\eta)\sin\theta_1}.$$

Using the fact that $K_1 I_1 \cong K_0 I_0$ for relatively large $\gamma_0 a$, $\sin^2\theta_1 \ll 1$ and the definition of x, the above equation may be simplified to

$$\frac{1}{2} < \frac{W}{W_I} \cong \frac{x^2}{x^2 + 2x - 1}. \tag{7.17}$$

W/W_I increases as $x = \csc\theta/\csc\theta_1$ increases. With x^2 as large as 6, i.e., with $\gamma_{0,0} a/\cot\theta$ less than 0.2, $W/W_I \cong 0.6$, being only slightly greater than $\frac{1}{2}$.

The above discussion is based on the one-term approximation which, we believe, should give reliable results. The characteristic properties are essentially determined by the symmetries of the structure rather than by the degree of approximation.

In Fig. 6 are plotted the energy densities of the first few components in single- and twin-helix structures having $\eta=1$, $ka=0.4$, $\beta_0 a=\beta_{0,0} a=1.888$, $\cot\theta_1=4.785$, and $\cot\theta=10$, respectively. W_0 is taken as unity and E_z on the axis is taken to be the same for both structures. We note that $W_{0,0} \cong 0.53 W_0$, $2W_{-1,0} \cong 0.50 W_{-1}$, and all other values of $2W_{l,0}$ are less than the corresponding values of W_l. The calculated value of W/W_I is about 0.59.

The energy flow or power P is obtained by multiplying the energy density by the group velocity

$$P = W \cdot v_g. \tag{7.18}$$

Both W and v_g are smaller for the twin-helix; the reduction in P is greater than that in either W or v_g. The two impedance parameters $Z_{0,0} = |E_z(r=a)_{0,0}|^2/\beta_{0,0}^2 P$ and $Z_0 = |E_z(r=a)_0|^2/\beta_0^2 P_I$ are inversely proportional to power, so $Z_{0,0}/Z_0 > W_I/W$. The penalty for a greater impedance parameter is, of course, in the reduction of the band width, though this reduction, as can be seen in Fig. 4, is only moderate.

For $l \neq 0$, $(E_z)_{l,0}$ and $(E_z)_l$ depend on the angle ϕ; $(E_z)_{l,0} \sim \cos l\phi$ and $(E_z)_l \sim e^{-jl\phi}$. We shall define the impedance parameters of space harmonics by averaging

$|(E_z)_{l,0}|^2$ over ϕ, i.e.,

$$Z_{l,0}=\frac{1}{\beta_{l,0}^2 P}\cdot\frac{1}{2\pi}\int_0^{2\pi} d\phi\,|E_z(r=a)_{l,0}|^2. \quad (7.19)$$

Z_l is defined like Z_0, i.e.,

$$Z_l=\frac{1}{\beta_l^2 P}|E_z(r=a)_l|^2. \quad (7.20)$$

$Z_{l,0}$ and Z_l of the first few space harmonics are compared in Fig. 7 for the same two cases as compared in the last figure. We note that $Z_{0,0}\cong 2.4Z_0$, $Z_{-1,0}\cong 0.16Z_{-1}$, and all other values of $Z_{l,0}$ are small compared to the corresponding values of Z_l.

The other pair of structures, $\cot\theta=5$, $\cot\theta_1=2.5$, $\eta=1$ have approximately the same relative characteristics, so need not be discussed in detail.

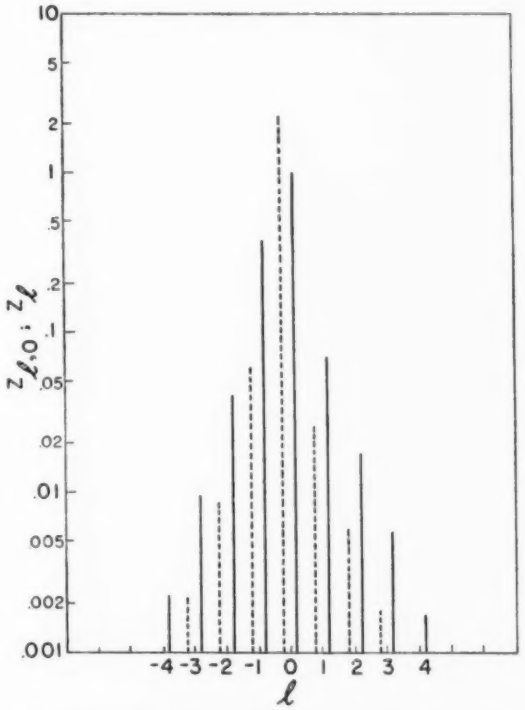


FIG. 7. Comparison of impedance parameters (see Eq. 7.19, 7.20) for the first few components of a single-helix ($\eta=1$, $\cot\theta_1=4.785$) and a twin-helix ($\eta=1$, $\cot\theta=10$). $ka=0.4$, $\beta_0 a=\beta_{0,0}a=1.888$. Arbitrary scale.

8. CONCLUSION

The detailed calculations made on the symmetric mode of a twin helix have borne out the qualitative predictions which were based on the symmetry of the structure. The principal result is that, for the symmetric mode, there is no *TE* portion of the field in the fundamental component. This means, numerically, that there is practically no magnetic energy stored in the fundamental component, and, therefore, the energy in the space harmonics has to be mainly magnetic. For the corresponding single helix of the same phase velocity and diameter, the energy density in the fundamental component is about equally divided between electric and magnetic, and the same holds for the space harmonics. As a consequence, the impedance of the fundamental component of the twin helix is about twice that for a single helix for the velocity and diameter ranges considered here. Conversely, the impedance of the space harmonics in the twin helix case is reduced by factors of the order of 5 to 10. This makes a twin helix much more desirable as a propagating structure for high-voltage traveling-wave tubes than the single helix.

Some measurements which have been made on phase velocity give quite good agreement with the theory, particularly for small η. These measurements were made with the helices in contact, and the departures from theory for large η may be due to the current approximations not being as good for wide helices in contact as was discussed in Section V on the distinction between insulated and noninsulated tapes. Measurements on an insulated twin-helix with $\eta=2$ did agree quite well with the calculation. Some cruder measurements on the impedance indicate that this is also in rough agreement with the theory. More details of these measurements will be reported later.

We wish to make acknowledgments to Mr. John R. Nevins, Jr., who made the measurements, and to Mr. Walter R. Ayres, who did most of the numerical work.

Evaporation of Barium and Strontium from Oxide-Coated Cathodes*

L. A. WOOTEN, A. E. RUEHLE,† AND GEORGE E. MOORE
Bell Telephone Laboratories, Inc., Murray Hill, New Jersey

(Received February 8, 1954)

Rates of evaporation of Ba and Sr from oxide cathodes in a practical tube structure were measured during life tests up to 20 000 hours. The effects on evaporation rates are shown for (a) impurity in the Ni support, (b) cathode temperature, and (c) space current. We conclude:

1. The product evaporating from commercial type cathodes under normal conditions is largely Ba metal; less than 5 percent is Sr; less than 2 percent is BaO; and less than 0.01 percent is SrO.
2. Chemical reducing agents in the support metal‡ are important in determining the rate of Ba evaporating during life.
3. A factor not closely controlled in the present experiments—anode and grid composition—affects the rate markedly. This is especially true for the evolution of Ba during exhaust.
4. Within the precision of these experiments, there is (a) no correlation between the rate of Ba evaporation and thermionic activity of individual cathodes, and (b) no effect of space current on the rate of Ba evaporation.

1. INTRODUCTION

MEASUREMENT of the evaporation rates of Ba, Sr, and their compounds from oxide cathodes is important for theories[1] relating chemical composition to thermionic behavior.

The objectives of our experiments were:

(A) To establish the amounts and nature of the material evaporating from oxide cathodes during life.

(B) To examine how the evaporation rate varies with: (1) operating temperature; (2) space current density; (3) composition of other electrodes.

(C) To seek a correlation between thermionic emission and quantity of evaporated Ba metal.

Spectrochemical tests of the product evaporated from a filamentary cathode were made in approximately 400 triodes having essentially the structure of a practical repeater tube. They were built especially for these tests and only one determination of the total evaporated product was possible for each tube. The spectrochemical method is sensitive and relatively free from interference by extraneous elements. However, its limitations led to additional experiments in which evaporation rates were measured from one filament under several conditions of operation. The tube resembled those employed by Becker and Sears[2] and used the thermionic activation of an adjacent W ribbon to measure Ba evaporated from the oxide-coated filament. Although this nondestructive method is even more sensitive than the spectrochemical, the two techniques are really complementary.

2.0 EXPERIMENTAL

2.1. Spectrochemical Tests

2.11. Tubes

The triode is shown diagrammatically in Fig. 1. The anode was bright 0.005 in. Ni (instead of carbonized as in the repeater tube), and the grid consisted of Mo laterals wound on Ni support wires. Mg getters and glass bead insulators replaced the usual Ba-Mg getter and mica insulator, for analytical reasons.

The Ni filament support was 0.95 cm² in area, weighed 13.5 mg, and was coated with a mixture of BaO and SrO containing 38.7 mole percent BaO. The "uncombined"[1] coating was applied as carbonates, 4 mg per cm², by the "drag" method. The filaments usually operated at a true temperature of 1020°±10°K. Temperatures were measured pyrometrically and corrected for the (lime glass) bulb and for spectral emissivity.[2]

During evacuation by a Hg diffusion pump, the glass was thoroughly baked at 430°C; the anodes were then heated by high-frequency induction, while the filaments were operated at approximately 1300°K for 2 minutes to decompose the carbonates to oxides. The getters were then flashed and the tubes sealed off.

Table I shows the compositions of the three supplies of nickel used as filament "core," or support metal. No. 1 and No. 2 Ni are typical of the so-called "Grade A" nickel supports used commercially, while No. 3 is much more pure.

2.12. Method of Spectrochemical Analysis

The anode and grid collected the product evaporated from the filament. To standardize effects caused by anode and grid materials, they were made in one large lot out of one homogeneous supply of material.§ The tubes were assembled, pumped, and aged in lots of

* Note: This program, on *filamentary* oxide cathodes, became inactive in December, 1941; however, questions raised during this program stimulated some of our post-war work. Although the results in paragraph 3 of the abstract are not yet understood in detail, the sudden death of L. A. Wooten influenced us to seek publication at this time. He led much of the experimental work but is not responsible for the interpretation.

† Present address: Mallinckrodt Chemical Works, St. Louis, Missouri.

‡ This phase of the work was discussed orally by L. A. Wooten, Phys. Rev. **69**, 248 (1946).

[1] J. P. Blewett, J. Appl. Phys. **10**, 668 (1939); and **17**, 643 (1946).

[2] J. A. Becker and R. W. Sears, Phys. Rev. **38**, 2193 (1931).

[3] G. E. Moore and H. W. Allison, J. Appl. Phys. **12**, 431 (1941).

§ The composition of the Ni anode was essentially the same as that of No. 1 Ni.

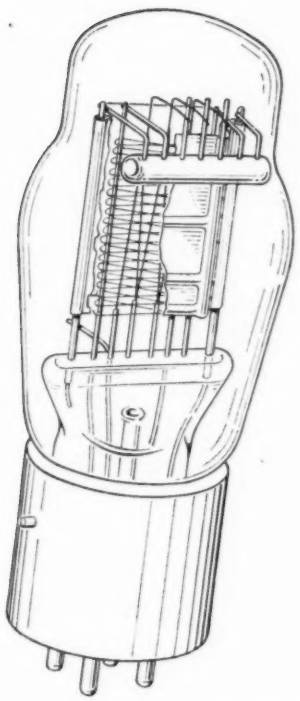


FIG. 1. Tube for spectrochemical tests.

40 to 100, employing identical procedure. Care was taken in assembling and handling to avoid contamination by mechanical transfer of coating. Immediately after exhaust, and at intervals during life, samples (3 to 15 tubes) were selected at random for analysis. Each tube was opened and dismantled carefully, and the anode and grid were stripped of cathode material by solution in warm dilute HCL, which was evaporated to 1 cc. A 0.1 cc aliquot was transferred to flat graphite electrodes of tested purity and analyzed spectrochemically by the comparison standard method. The high-voltage alternating-current arc[4] was used for excitation.

With a Hilger E-1 spectrograph and Eastman DC Ortho plates, sensitivity was better than 0.3 microgram Ba per cc of solution at λ4934.1 A. The excitation method gave low background at the cyanogen band so that λ4077.7 A gave a simultaneous sensitivity of 1 microgram of Sr per cc. Photographic densities were measured by a modified Moll densitometer.

The precision for Ba was about ±7.5 percent determined by tests on synthetic solutions of known Ba content.[4] Analogous tests gave a precision of about ±20 percent for Sr. Repeated tests indicated equal precision for samples from tubes.

Error might occur in measurement of the amounts evaporated by:

(a) Contamination of anode or grid by mechanical transfer of coating. Virtual absence of Sr on the collectors was satistactory evidence against contamination, because coating mechanically transferred contained nearly equal parts of Ba and Sr. Seven tubes were so rejected; these comprised less than 2 percent of the total tested.

(b) Loss of evaporated material by scattering or reevaporation to remote parts of the tube. Bulb walls and support wires were usually analyzed and in no case did the Ba and Sr detected exceed 5 percent of the total found in the entire tube. While these tests show that scattering and re-evaporation are experimentally unimportant, reevaporation is further discussed in Appendix I.

TABLE I. Composition of nickel support.

	Ni No. 1 wt percent	Ni No. 2 wt percent	Ni No. 3 wt percent
Magnesium	0.09	0.12	0.005
Silicon	0.02	0.08	0.01
Carbon	0.03	0.03	...
Iron	0.24	0.21	0.01
Manganese	0.25	0.23	<0.001
Copper	0.14	0.11	0.005
Cobalt	0.05	0.05	0.001
Aluminum	0.001	0.001	<0.001
Chromium	0.01	0.01	<0.001
Silver	0.001	0.001	0.001
Nickel (by difference)	99.17	99.16	99.97

[4] A. E. Ruehle and E. K. Jaycox, J. Ind. Eng. Chem. (Anal. Ed.) 12, 260 (1940).

2.2. Thermionic Tests

2.21. Experimental Tube

Evaporation rates were measured thermionically in a tube placed at our disposal by R. W. Sears. This is shown schematically in Fig. 2. Filaments F_1 and F_2 were essentially the same as those using No. 1 support in the spectrochemical tests, and W was a tungsten ribbon 0.050 cm wide by 0.0025 cm thick. It could be cleaned whenever desired by heating to 2700°K for a few seconds; it intercepted material evaporated from F_1. Becker[2] and others have shown that the presence of a monolayer of Ba on such W ribbons gives an optimum thermionic emission. S was a shield of Ni sheet movable in its own plane. All filaments were 15 cm long; emission

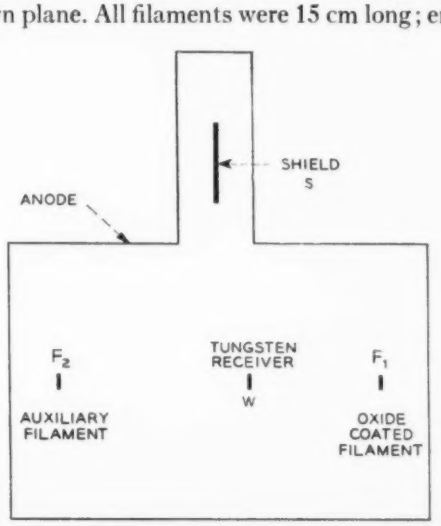


FIG. 2. Tube for thermionic tests.

from the central 2 cm was measured to a guard-ring-anode system.

2.22. Method of Thermionic Analysis

This technique has been described.[2,5] Its great advantage is the fact that a monolayer (0.11 μ gram per cm^2) can be measured repeatedly with a precision of 5 percent without destruction of the system; thus the evaporated product can be investigated under many operating conditions. When, and only when, space current was drawn from F_1 of Fig. 2 to the anode, a deactivating gas (probably CO or CO_2) was deposited on W, together with the activating agent. The deactivating gas was removed from W by bombardment with electrons from F_2. These were drawn around the shield 3 which was placed between F_2 and F_1-W filaments to prevent material evaporated by F_2 from depositing on either F_1 or W. The deactivating component thus disappeared from W, leaving the activating agents such as Ba, which were unaffected by additional electron bombardment.

2.3. Limitations of Methods

The evaporated product is a mixture of Ba and Sr metals with their oxides. Although the spectrochemical method distinguishes between total Ba and total Sr, it gives no information on the oxide-to-metal ratio for either Ba or Sr.

The thermionic activation of W ribbons by a film of pure BaO (or SrO) is more pronounced[5] than by a film of the corresponding metal; however, the method gives very little information on the composition of mixtures. Therefore, the oxide-to-metal ratio for either Ba or Sr must be inferred somewhat indirectly, in either type of experiment.

3. RESULTS AND DISCUSSION

3.1. Amounts and Nature of Evaporated Material[||]

3.11. General Spectrochemical

Tables II and III show results for four groups of tubes using No. 1 and No. 2 Ni, respectively, while Table IV shows the product evaporated from filaments of No. 3 Ni. Some of these results are also shown in Fig. 3. These groups were aged with normal anode potential (135 volts) giving a space current of approximately 10 ma per cm^2.

The second column gives the number of tubes used for one experimental point. (Two or more determinations were made on each tube.) The total amounts of Ba and Sr (γ) and the standard deviations, σ, are expressed in micrograms.

Within the relatively large σ shown in Tables II to IV (and in Fig. 5), the behavior of three different groups containing No. 1 Ni and aged with space current were generally reproducible. The three most significant results of Fig. 3 are:

(a) The spectrochemical data show that the evaporated product is nearly all Ba or BaO, with not over 5 percent Sr or SrO.

(b) The two stages in life for Nos. 1 and 2 Ni characterized by different rates of Ba production. During the first 1500 hours approximately, the evaporation rate of total Ba is 30 to 50 μg per cm^2 per 1000 hours. After additional aging, the rate falls to about 2 μg per 1000 hours.

(c) For No. 3 Ni, the initial stage of rapid evaporation does not occur. From these measurements, one might conclude that essentially nothing evaporates after exhaust, if the filament core is of sufficiently pure Ni. This is strong evidence that reduction of BaO by impurities such as Mg, C, and Si in Nos. 1 and 2 Ni, causes a high rate of Ba metal production. In confirmation, spectrochemical analysis shows that Si compounds and MgO accumulate at the interface of aged filaments using No. 1 and No. 2 nickel support metal.

TABLE II. Evaporation data on No. 1 Ni filament.

	Age (hrs)	No. of tubes	Ba (γ)	Deviation (σ)	Sr (γ)	Deviation (σ)
Group 1	0	16	47	±10.5	<5	...
	1060	7	88	7.2	<5	...
	2165	7	106	19.2	<5	...
	3175	6	105	16.5	<5	...
	5130	6	109	10.6	6.0	2.0
	10075	5	114	2.4	4.0	4.1
	20000	5	137	12.0	8.0	5.3
Group 2	0	4	52	7.0	2.5	±1.1
	156	4	75	16.9	2.5	1.1
	388	4	85	15.4	4.2	0.4
	640	4	94	7.2	2.8	0.4
	990	6	102	14.3	3.7	1.4
	1300	4	100	9.4	2.1	1.7
	1615	4	104	16.3	2.0	0.0
	2015	4	106	6.5	2.0	0.7
	2490	4	108	13.0	3.0	1.2
	4020	3	112	2.4	...	...
	0	15	53	18.4	1.2	0.3
	500	5	73	9.0	2.4	1.4
	1000	5	105	23.6	2.0	0.7
	2000	4	95	11.4	2.9	0.9
	7000	5	115	22.2	5.7	3.3

[a] Analyzed by the dc-arc method. (See reference 4.)

3.12 Information from Other Experiments

These results, together with three other experiments now quoted, lead to the conclusion that nearly all of

TABLE III. Evaporation data on No. 2 Ni filament.

Age (hrs)	No. of tubes	Ba evaporated (γ)	(σ)	Sr evaporated (γ)	(σ)
0	6	78	±17.9	3.0	2.1
640	1	100	...	...	0
990	5	102	19.9	6.0	2.9
2015	3	107	12.5	5.0	2.1
4022	3	108	2.5	...	...

[5] G. E. Moore and H. W. Allison, Phys. Rev. **77**, 246 (1950).

[||] Similar work, done since the war, on the effect of impurities (in the nickel of equipotential cathodes) was discussed in a preliminary way by H. E. Kern; for abstracts see Phys. Rev. **78**, 355 (1950), and Phys. Rev. **82**, 574 (1951).

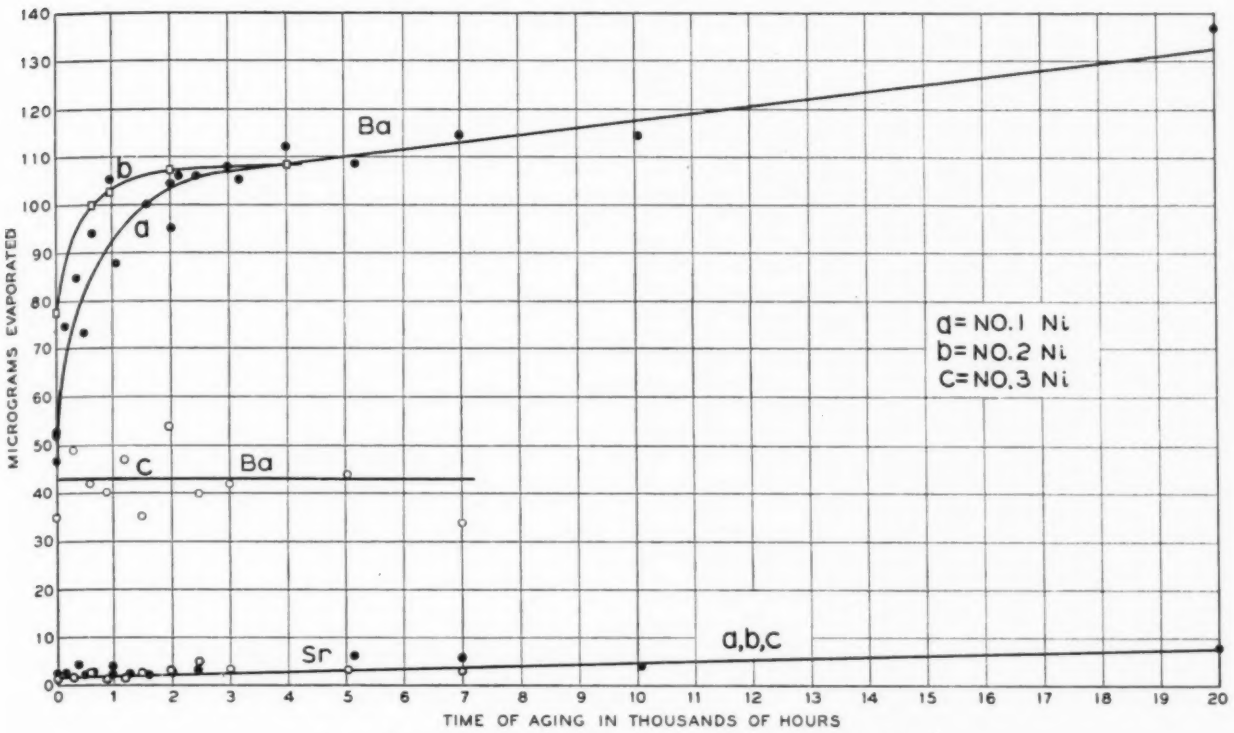


FIG. 3. Effect of nickel composition on evaporation of Ba and Sr.

TABLE IV. Evaporation data on No. 3 Ni filament.

Age (hrs)	No. of tubes	Ba evaporated (γ)	(σ)	Sr evaporated (γ)	(σ)
0	15	35	±11.1	1.3	±0.4
300	5	49	6.5	1.6	0.2
600	5	42	9.3	2.5	0.5
900	5	40	6.8	1.7	0.4
1200	5	47	8.6	1.3	0.2
1500	7	35	12.8	2.4	1.2
2000	4	54	11.1	3.0	1.1
2500	5	40	9.6	4.8	1.8
3000	5	42	8.9	3.2	1.4
5025	5	44	8.5	3.2	1.2
7015	6	34	18.0	3.0	2.1

the total Ba and total Sr evolved leave the filaments as free metal.

3.121. Vapor pressure measurements.—For the deposits found immediately after exhaust (even including filaments of No. 3 Ni), vapor pressure measurements of BaO[6,7] show that less than 2 percent of the total Ba could have evaporated, at 1300°K, as BaO. From similar measurements [6,8] of SrO, less than 0.01 percent of the total Sr found could have evaporated as SrO.

For the rate of evaporation from filaments of No. 1 and 2 Ni during the first 2000 hours, vapor pressures suggest that the product is nearly all Ba and Sr metal. After about 2000 hours the spectrochemically indicated rate of total Ba evolution (2 μg per 1000 hours) is only about half the evaporation rate of pure BaO. However, this measurement could still imply nearly pure Ba metal, because, from the solid solution always formed[9,10] by (BaSr)O, the rate of BaO evaporation quickly decays as a result of depletion in the surface layers. Our own spectrochemical data on sections through the coating also indicate that BaO is strongly depleted in the surface layers of aged coatings. Electron diffraction studies[11,12] indicate the complete absence of BaO on the surface of aged cathodes of mixed oxides.

3.122. Tests by Evolution of H_2 in Contact with H_2O.—Interpretation was also confirmed by many tests¶ in which the Ba *metal* only, was measured. Coated filaments of No. 1 and No. 2 nickel were mounted in bulbs with no other electrodes. They were then aged at 1020°K (without space current, of course) and connected to a gas analysis system without loss of vacuum, by using a break-off tip. H_2O vapor was then admitted, and allowed to react with the Ba and Sr metals previously produced in the tube. At ambient temperature, this produced H_2 gas which was a measure of the free Ba or Sr metal;[13–15] BaO will produce no H_2. The rate of Ba metal evolution after exhaust agreed with the initial rate shown in Fig. 3, but prolonged tests of this kind are not practical.

[6] A. Claassen and C. F. Veenemans, Z. Physik **80**, 342 (1933).
[7] Blewett, Liebhafsky, and Hennelly, J. Chem. Phys. **7**, 478 (1939).
[8] Moore, Allison, and Struthers, J. Chem. Phys. **18**, 1572 (1950).
[9] M. Benjamin and H. P. Rooksby, Phil. Mag. **15**, 810 (1933).
[10] W. G. Burgers, Z. Physik. **80**, 352 (1933).
[11] H. Gaertner, Phil. Mag. **19**, 82 (1935).
[12] J. A. Darbyshire, Proc. Phys. Soc. (London) **50**, 635 (1938).
¶ To be submitted to J.Appl. Phys.
[13] L. A. Wooten, ASTM Bull. **108**, 39 (1941).
[14] G. E. Moore, J. Chem. Phys. **9**, 427 (1941).
[15] Moore, Allison, and Morrison, J. Chem. Phys. **18**, 1579 (1950).

3.123. Tests by thermionic method.—Up to 7000 hours, when tests were discontinued, the thermionic method affords the following evidence that the product from filaments using No. 1 or 2 Ni is nearly all Ba metal:

(a) The optimum observed emission of the *W* ribbon corresponds to that obtained with films known to be pure Ba metal.

(b) On heating the *W* ribbon coated with evaporated product, its emission decay resembles that found for *W* ribbons known to be coated with pure Ba metal, rather than BaO.

(c) When the *W* ribbon is heated, about 1/20 of the product on its surface reevaporates to the oxide cathode whence it originally came. Considerably less than this fraction of a monolayer increases the emission of the oxide cathode (measured at a low temperature) by a factor of 50, which is not altered by additional reevaporated product. This activation is analogous to that obtained from sources known to produce Ba metal, and is not stable for even 1 minute at 1020°K.

This evidence indicates that nearly all the material evaporated from filaments of Nos. 1 or 2 Ni during the first 7000 hours (and probably much longer) is Ba metal. The decrease in Ba evaporation rate noted after about 2000 hours is probably caused by depletion of impurities near the interface, subsequent reaction being limited by diffusion of reducing agents. Of the total Sr evaporated, essentially all is Sr metal. From filaments using the purer No. 3 Ni, the amounts are so small that analysis by either method is difficult.

3.2. Effect on Evaporation Rate of

3.21. Temperature.

For Tables II, III, and IV, and Fig. 3 the corrected filament temperature was 1020°K. Additional tubes

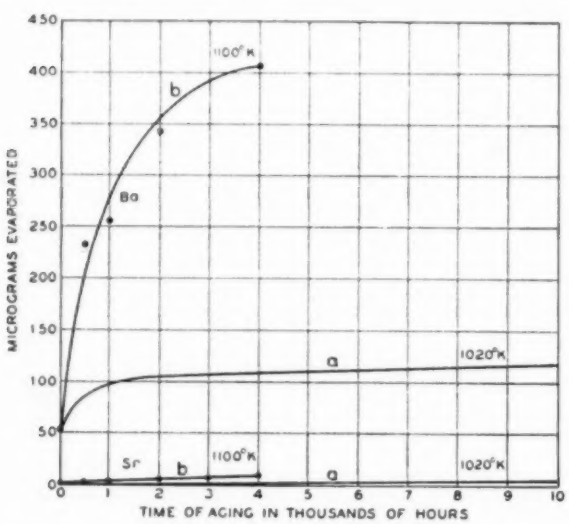


FIG. 4. Effect of temperature on evaporation of Ba and Sr. (a) 1020°K (Curve *a* of Fig. 3 reproduced without points). (b) 1100°K.

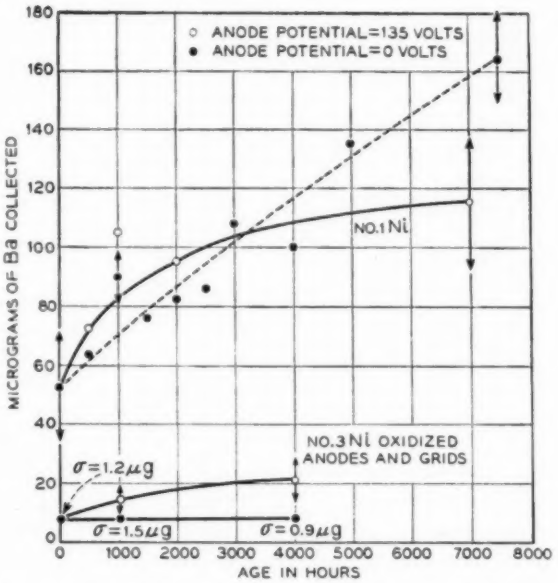


FIG. 5. Effect of space current on evaporation of Ba. (a) Upper curves—apparent effect of space current. Filaments of No. 1 Ni. (b) Lower curves—apparent effect of space current. Filaments of No. 3 Ni, oxidized Ni anodes and grids.

from the same lot as Group 3 in Table II, also with filaments of No. 1 Ni, were aged at a corrected temperature of 1100°K and 135 volts E_b. Figure 4 shows the effect; curve a of Fig. 3 is reproduced for comparison. An increase of 80°K produces an approximately 4-fold increase in the total amount of Ba evaporating between zero and 4000 hours. Some increase is also shown for Sr evaporation. This sensitivity, together with observed temperature variations (about ±10°K) under normal operating conditions, could partially explain the large magnitudes of σ.

3.22. Effect on Evaporation Rate of: Space Current Density

3.221. Spectrochemical tests.—In one theory of thermionic activation,[2,16] space current is considered important for production of excess Ba. To evaluate its effect, three groups from the same lots as above, and No. 1 Ni filament, were aged at 1020°K and without space current ($E_B=0$). The two upper curves of Fig. 5 show one comparison—one group with, and one group without space current. The standard deviation, σ, is shown for some typical points. Although σ is too large for definite judgment of the effect of space current, one might conclude from the curves that electron flow increases Ba evolution for the first 2000 to 3000 hours and decreases it thereafter. Similar indications were obtained in two other statistical comparisons of this kind.

[16] (a) H. Rothe, Z. Physik **36**, 737 (1926); (b) W. Espe, Wiss. Veröffentl. Siemens-Konzern **5**, 29 (1927); (c) F. Detels, Jahrb. d. Drahtl. Telegr. u. Teleph. **30**, 10–14, 52–59 (1927); (d) J. A. Becker, Phys. Rev. **34**, 1323 (1929); Trans. Electro-Chem. Soc. **59**, 207 (1931).

3.222. Thermionic tests.—For more sensitive non-statistical evaluation of the effect of space current, thermionic tests were made in the tube of Fig. 2 at space currents of 0, 10, and 20 ma per cm^2 over a 7000-hour life period. Typical results are shown in Fig. 6. The temperature required in the *W* ribbon for a fixed emission density of 4.5×10^{-8} amp/cm^2 is plotted[5] against the time for which the oxide-coated filament was operated at 1030°K. These curves always begin at 1470±10°K, appropriate for clean *W*, and gradually decrease to a minimum of 715°K, appropriate for a monolayer of pure Ba metal on *W*. The receiver then becomes less active with additional deposit. The points show that the activating agent arrives at the *W* ribbon at the same rate for any space-current density from 0 to 20 ma per cm^2. This was duplicated in repeated tests at these varied conditions.

Several tubes with equipotential cathodes had previously been studied by R. W. Sears** and others, using the thermionic method. In each, deposit of the activating agent occurred at the same rate, with or without space current. Some of these cathodes were coated only with SrO.

We ignore spectrochemical results shown in Fig. 5, and conclude from the thermionic data of Fig. 6 that Ba metal evolution is independent of space current within experimental precision. If space current itself has any effect on Ba production and evaporation rates, this effect is very minute compared to that of impurities in the support metal or anode material.

However, major improvement in emission is always noted in these filaments when space current flows, and *might* be caused by concentration[17] of excess Ba at favorable locations in the coat. This implies that excess Ba is important for thermionic emission at practical temperatures, an assumption which has never really been verified.

3.23. Effect on Evaporation Rate of: Composition of Other Electrodes

3.231. Results Implicit in Preceding.—Three results, not yet emphasized, can probably be explained by interaction between filament coating and impurities from the anode and grid.

(a) As stated, the amounts of total Ba and Sr collected during exhaust are far too large to have evaporated as oxides; however, the tests by H_2 evolution, using filaments without anode or grid, indicate that these amounts are also too large, by a factor of 5 to 10, to have been produced as free Ba and Sr metal by reaction with impurities in the Ni support.

(b) If all impurities present in Nos. 1 and 2 Ni reacted with BaO to produce Ba metal, the Ba produced would have been considerably less than the amounts measured and listed in Tables II and III. The impurity in No. 3 Ni is not sufficient to account even for the Ba evolved during exhaust, shown in Fig. 3.

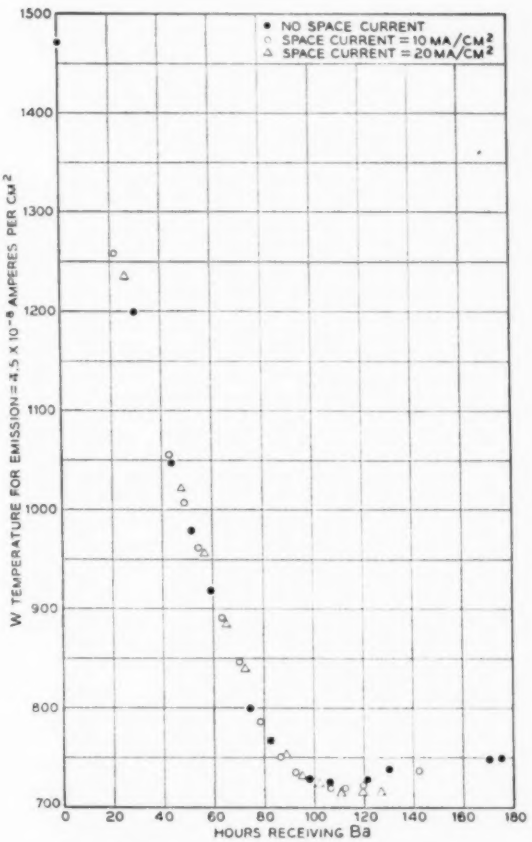


FIG. 6. Effect of space current—thermionic method.

(c) The large values of σ, present following exhaust, change slightly or not at all during aging. Although this could be the result of uncontrolled variations of temperature (±15°) during exhaust, the variations in σ are not well correlated with measured variations in temperature.

3.232. Experiments with Oxidized Ni as Anode and Grid.—To assist in interpretation of these results, 40 additional tubes were prepared with anodes and grids (including laterals) of oxidized Ni. The subassemblies were first cleaned in H_2 for 1 hour at 975°C and then heated in purified O_2 for 1 hour at 975°C, giving a surface film of NiO of weight 0.9 to 1.4 percent that of the 0.005 in. Ni sheet. No. 3 Ni filaments were used. Half the tubes were aged with anode voltage and half without. The results, shown in the lower curves of Fig. 5 suggest that space current increases Ba production. More striking is the much lower initial point than for the same filament in bright Ni anodes, shown in Fig. 3. Furthermore, σ for these tubes (indicated in Fig. 5) was negligible after exhaust and remained negligible in those tubes aged without space current, but increased markedly when space current flowed. This

** Private Report.

[17] J. E. Harris, in memoranda circulated in these laboratories, has reported that Ba was present in the Pt base of filaments aged with space current, and was not present when no space current was drawn. See also reference 6 (d).

latter is not understood in detail. In an earlier test with parts less carefully oxidized, all tubes were operated with space current; the total Ba evolution was considerably higher.

With oxidized electrodes the measured impurity content of No. 3 Ni is almost sufficient to account for the total Ba deposit immediately following exhaust; some of this left the filament as BaO, of course.

3.233. Mg in outer layers of coat.—The influence of volatile reducing agent in the anode, etc., is also indicated after exhaust by Mg, detected spectrochemically, in the outer layers of the (Ba Sr)O coat in any tubes with Nos. 1, 2, or 3 Ni, using bright Ni anodes. A number of tubes aged without getters showed that the Mg getter did not affect the initial point, σ, or the rate of Ba production during life. From these results, we conclude that the composition of other electrodes is important for determining the evolution of Ba and Sr metal from the cathode. To determine mechanisms will require additional work.

3.3. Do Thermionic Emission and Evaporated Ba Metal Correlate?

Whenever tubes were selected for chemical analysis, "total emission" of every filament in the entire group was tested; these thermionic emissions always covered a wide range. Emission was measured by the ballistic deflection of a suitable galvanometer on the grid-anode terminal, when a 4 mfd condenser, charged to 400 volts, was discharged through the tube connected as a diode. Emissions up to about 6 amperes in this structure could have been measured in this way.

Immediately after exhaust, the emission of all filaments in the present experiments was low; this always improved on aging. Many attempts were made to correlate emission with the amount of Ba or Sr observed on the collectors. Two such tests are shown as Fig. 7(a) and 7(b). Both are for experiments with No. 3 Ni, shown in curve *c* of Fig. 3. Total emission is plotted against total Ba and total Sr collection. The first chart gives results on the 15 tubes tested immediately after exhaust, and the second is for the 5 tubes removed from life for spectrochemical test at 600 hours. Although the mean emission improved during this 600-hour interval by a factor of 780, statistics in Table IV do not indicate clearly that any additional Ba had left the filament. In common with all other tests throughout these experiments, no correlation was found.

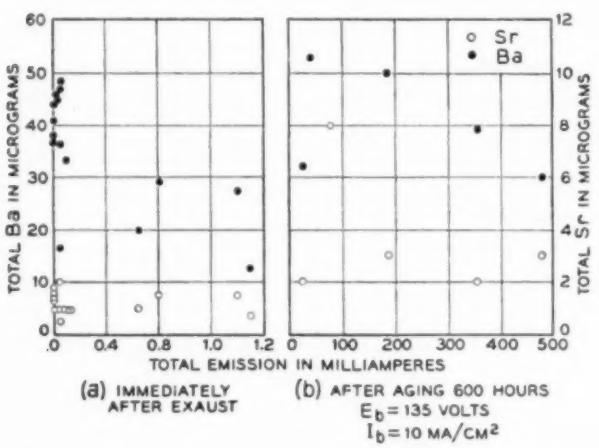


FIG. 7. Attempted correlation of emission with Ba or Sr found on electrodes. Note change in emission scale between parts (a) and (b).

The mean emission of commercial repeater filaments of this type is usually 1.5 amperes after preliminary aging. This declines by a factor of about 2 during 15 000 hours, and more slowly thereafter. With the exception of those with oxidized anodes and grids and those aged at higher temperature, the mean emission of all filaments of No. 1 or No. 2 Ni in these experiments, after preliminary aging, was adequate for repeater use (150 ma) but averaged not over $\frac{1}{4}$ that usually observed in actual repeater tubes. In particular, the emission of all 5 tubes remaining on test for 10 000 hours, was adequate.†† During early life, the mean emission of filaments of No. 3 Ni was considerably less than for those employing commercial Ni, such as No. 1 or No. 2. The emission of the filament used in the thermionic method was about $\frac{1}{2}$ that found in repeater tubes. As generally observed, the emission during space-current aging in any of these tubes is about 4 times that without space current. Emission in the presence of Mg getters was usually several fold larger than in the absence of getters, but Ba evolution was unaffected. These wide differences in emission are obtained with no measurable difference in Ba evolution rate.

Because these experiments show no correlation between rate of Ba evolution and thermionic emission, they provide no support for the theory that high thermionic activity in oxide cathodes depends on excess Ba metal. However, the results do not warrant the opposite conclusion that the Ba produced is just a useless by-product of chemical processes in the cathode system. Perhaps the safest conclusions are that the work (1) illustrates various factors affecting the rate of Ba production by chemical reactions, (2) shows that space current has essentially no effect on Ba evolution, (3) apart from the more rapid activation of Nos. 1 and 2 Ni filaments than for the No. 3 Ni, there is no *clear* correlation between the speed of the chemical processes and the electron emission of the filaments at any stage of life.

3.4. Relation to Published Work on Oxide Cathodes

Theories of the oxide cathode are well summarized by Nergaard in a paper[18] which also proposes a

†† The mean life of repeater tubes of this general type is of the order of 10^5 hours. One group of 216, made during development of this tube, has been on continuous life for over 16 years without a single failure. Half were operated at 20 ma per cm² and half at 8 ma per cm².

[18] L. S. Nergaard, RCA Rev. 13, 464 (1952).

new and ingenious theory. In undertaking our work, primary emphasis was on the measurement of a phenomenon, the evolution of Ba metal, which one might expect to correlate with emission, according to any published theory. The fact that no correlation was observed does not necessarily argue against any of these theories. For example, on Nergaard's depletion layer hypothesis, one might expect space-current flow to decrease the rate of Ba evolution, especially from cathodes of lower thermionic activity. However, the theory is necessarily so complex, that our result is not in definite contradiction.

The fact that Ba evolution was independent of space-current flow (in the above tests and many others not described) may seem inconsistent with frequent reports that space current causes oxygen evolution. Published reports on oxygen evolution are rather contradictory. Thus Barton,[19] Bachman and Carnahan,[20] Broadway and Pearce,[21] and Sloane and Watt[22] all observe O_2^-, and some report O^- while Becker[16d] and Plumlee and Smith[23] state that the oxygen is uncharged. We have made preliminary measurements by mass spectrometer of materials evolved from oxide cathodes. On the few occasions when oxygen peaks were visible, we found no evidence that the gas came directly from the cathode.

Our results are in general agreement with those presented by Leverton and Shepherd,[24] but they differ in two respects:

(1) During exhaust, they found essentially equal amounts of Ba and Sr transferred, although the Sr component decreased when their cathodes were aged under typical, or accelerated, life conditions. Their cathodes were heated about 100° higher during exhaust than in our work.

(2) They found a 50-fold reduction in total Sr transfer when space current flowed; they attribute this to stripping the Sr from the anode by electron bombardment. They did not observe the effect on Ba. In our experiments neither the reduction nor any stripping of either Sr or Ba was observed; they would have been detected, had they occurred.

Where our results appear to disagree with those published by others, we believe that both measurements represent accurate tests of systems which differ physically and chemically. This belief led to publication of detailed chemical analyses for the filament metal in the present work.

[19] H. A. Barton, Phys. Rev. **26**, 360 (1925).
[20] C. H. Bachman and C. W. Carnahan, Proc. Inst. Radio Engrs. **26**, 529 (1938).
[21] L. F. Broadway and A. F. Pearce, Proc. Phys. Soc. (London) **51**, 335 (1939).
[22] R. H. Sloane and C. S. Watt, Proc. Phys. Soc. (London) **61**, 217 (1948).
[23] R. H. Plumlee and L. P. Smith, J. Appl. Phys. **21**, 811 (1950).
[24] W. F. Leverton and W. G. Shepherd, J. Appl. Phys. **23**, 787 (1952).

ACKNOWLEDGMENT

We are indebted to J. R. Wilson and R. W. Sears for discussion throughout this work, and to V. L. Ronci and his colleagues for care in constructing the experimental tubes. The cooperation of H. W. Allison is also gratefully acknowledged.

APPENDIX I

Most of the energy dissipated in the anode is radiated; the resulting anode temperature is about 600°K when space current flows, and about 520°K without space current. The amounts of Ba present (100 micrograms) are equivalent to about 40 monolayers and at the probable temperature, about 0.02 micrograms would reevaporate[25] per hour per cm^2, if the Ba were present as bulk metal. The rate would be much more rapid during exhaust at anode temperatures of 1050°K. The anode area is about 20 cm^2. Because of the comparative absence of Ba from the bulb, etc., we may be sure that no such re-evaporation actually occurred from the anode. This indicates that Ba and Sr did not remain as a pure metal phase after deposition on the Nickel. Possible reasons for this are:

(a) Oxidation of Ni and Mg on the surface of the anode and grid by CO_2, etc., given off during exhaust. These oxides could hold Ba and Sr more tenaciously than clean Ni.

(b) As Ba and Sr are deposited, they could be oxidized by CO_2 and CO. Mass spectrometer tests[26] show these gases to be very generally present in such tubes.

In any precise spectrochemical determination of evaporation rates beyond 2000 hours, precautions not employed in the present work will be necessary to avoid re-evaporation effects.

[25] E. Rudberg and J. Lempert, J. Chem. Phys. **3**, 627 (1935).
[26] See reference 23—measurements made here support their result on the ubiquity of CO and CO_2.

Effect of Delayed Fission Neutrons on Reactor Kinetics

LAWRENCE BAYLOR ROBINSON*
Radiation Division, Naval Research Laboratory, Washington D. C.
(Received March 15 1954)

A detailed treatment of reactor kinetics is given. The kinetic equations are derived from a diffusion-Fermi age model. An iterative technique and a matrix method are discussed as possibilities for obtaining solutions. The relationship between the solutions with and without delayed neutrons is pointed out. The concept of stability is discussed for the system which includes the delayed neutrons. The similarity of the treatment here to that used in other branches of mathematical physics (especially quantum mechanics) is noted.

I. INTRODUCTION

AN adequate treatment of many aspects of reactor kinetics has been given by Chernick.[1] This treatment has been extended in regard to the concept of reactor stability.[2] In this report, the subject of reactor kinetics will be treated in a more detailed fashion so that the effect of delayed neutrons may be considered. Full use will be made of existing mathematical concepts and theorems which have been developed relating to the solution of differential equations. The applicability of such theorems in obtaining practical solutions will be pointed out. The emphasis will be on the techniques; numerical results will not be given.

II. DERIVATION OF EQUATIONS

The model used will be that of the one group thermal neutron-diffusion-Fermi age approximation.

The appropriate equations are[3]

$$L^2\nabla^2\phi(\mathbf{r},t)+[(1-\beta)k(t)\exp(-B^2t)-1]\phi(\mathbf{r},t)+\frac{p\exp(-B^2\tau)}{\sum_a}\sum_{i=1}^{m}\lambda_iC_i(\mathbf{r},t)=l_0\frac{\partial}{\partial t}\phi(\mathbf{r},t), \quad (1)$$

$$\frac{\partial}{\partial t}C_i(\mathbf{r},t)=-\lambda_iC_i(\mathbf{r},t)+\frac{k(t)}{p}\beta_i\sum_a\phi(\mathbf{r},t). \quad (i=1, 2, \cdots, m). \quad (2)$$

The symbols used above have the following meanings:

ϕ = thermal neutron flux (neutrons cm^{-2} sec^{-1}),
L = thermal diffusion length (cm),
B^2 = critical buckling (cm^{-2}),
τ = Fermi age (cm^2),
$\sum_a$ = over-all macroscopic absorption cross section (cm^{-1}),
$k(t)$ = infinite multiplication constant (time dependent),
C_i = concentration of precursor of ith delayed neutron group (atoms cm^{-3}),
λ_i = decay constant of C_i (sec^{-1}),
β_i = fraction of total fission neutrons in ith delayed group,
β = sum of the various β_i,
p = resonance escape probability,
l_0 = mean neutron lifetime in infinite medium (sec),
$\sum_{i=1}^{m}$ = means summation over the m groups of delayed neutrons.

The derivation is essentially that given by Glasstone and Edlund; in the present case k may be time dependent.

It will be assumed that the reactor is operating in an equilibrium state and that a small arbitrary change occurs in k displacing the dynamical system from equilibrium; the spatial distribution of the flux will be taken as remaining the same. It will also be assumed that the concentration of each delayed-neutron precursor is everywhere proportional to the space part of the flux. In other words, if one writes

$$\phi(\mathbf{r},t)=\phi(\mathbf{r})n(t), \quad (3)$$

$$C_i(\mathbf{r},t)=C_i(\mathbf{r})H_i(t), \quad (4)$$

then $\phi(\mathbf{r})/C_i(\mathbf{r})$ is independent of position. When Eqs. (3) and (4) are substituted in Eq. (2), one obtains

$$\frac{d}{dt}H_i(t)=-\lambda_iH_i(t)+\frac{k(t)}{p}\beta_i\sum_a\frac{\phi(\mathbf{r})}{C_i(\mathbf{r})}n(t). \quad (5)$$

($i=1, 2, \cdots, m$, where m is the number of groups of delayed neutrons.) The same substitutions in Eq. (1) yield

$$L^2\frac{\nabla^2\phi(\mathbf{r})}{\phi(\mathbf{r})}+[(1-\beta)k(t)\exp(-B^2\tau)-1]+\frac{p\exp(-B^2\tau)}{\sum_a}\sum_{i=1}^{m}\frac{C_i(\mathbf{r})H_i(t)}{\phi(\mathbf{r})n(t)}=\frac{l_0}{n(t)}\frac{d}{dt}n(t). \quad (6)$$

Remembering that the equilibrium space distribution of the flux obeys the equation,

$$\nabla^2\phi(\mathbf{r})+B^2\phi(\mathbf{r})=0, \quad (7)$$

* Present address: Department of Physics, Brooklyn College, Brooklyn, New York.

[1] J. Chernick, *The Dependence of Reactor Kinetics on Temperature* (Brookhaven National Laboratory report, BNL-173, December 20, 1951).

[2] L. B. Robinson, J. Appl. Phys. **25**, 516 (1954).

[3] S. Glasstone and M. C. Edlund, *The Elements of Nuclear Reactor Theory* (D. van Nostrand Company, Inc., New York, 1952), pp. 295–297.

one obtains

$$\frac{dn(t)}{dt}=\left[\frac{k(t)(1-\beta)\exp(-B^2\tau)}{1+L^2B^2}-1\right]\times\frac{n(t)}{l}+\frac{p\exp(-B^2\tau)}{l(1+L^2B^2)\sum_a}\times\left(\lambda_1\frac{C_1(\mathbf{r})}{\phi(\mathbf{r})}H_1(t)+\cdots+\lambda_m\frac{C_m(\mathbf{r})}{\phi(\mathbf{r})}H_m(t)\right), \quad (8)$$

where $l=l_0(1+L^2B^2)^{-1}$. It is clear that one does not need to restrict the treatment to constant values of k (i.e., a step increase) since the reduction of the partial differential equation to ordinary differential equations is independent of the time dependence of k.

In order to put in display (in a more clear fashion) the type of system of differential equations involved in this study, Eq. (5) will be written out in more detail.

$$\begin{aligned}
\frac{dH_1(t)}{dt}&=-\lambda_1H_1(t)+\frac{k(t)}{p}\beta_1\sum_a\frac{\phi(\mathbf{r})}{C_1(\mathbf{r})}n(t),\\
\frac{dH_2(t)}{dt}&=-\lambda_2H_2(t)+\frac{k(t)}{p}\beta_2\sum_a\frac{\phi(\mathbf{r})}{C_2(\mathbf{r})}n(t), \quad (5)\\
&\cdots\cdots\cdots\cdots\\
\frac{dH_m(t)}{dt}&=-\lambda_mH_m(t)+\frac{k(t)}{p}\beta_m\sum_a\frac{\phi(\mathbf{r})}{C_m(\mathbf{r})}n(t).
\end{aligned}$$

The equations relating the temperature T of the reactor, $k(t)$, and the temperature coefficient of multiplication (i.e., $dk/dT=\alpha$) are

$$\left[\frac{k(t)(1-\beta)\exp(-B^2\tau)}{1+L^2B^2}-1\right]=k_{ex}(t)=k_{ex}{}^c(t)-\alpha T, \quad (9)$$

$$\frac{d}{dt}T(t)=\gamma n(t)-\lambda T(t), \quad (10)$$

where $k_{ex}{}^c(t)$ is the induced change in excess multiplication k_{ex}, λ is the relaxation constant for heat removal and γ is a conversion constant involving heat released by fission and the specific heat of the medium. When Eq. (9) is substituted in Eq. (8), one obtains

$$\frac{d}{dt}n(t)=\frac{k_{ex}(t)n(t)}{l}+\Lambda(\mu_1H_1(t)+\cdots+\mu_mH_m(t)), \quad (11)$$

where

$$\Lambda=p\exp(-B^2\tau)l^{-1}(1+L^2B^2)^{-1}\sum_a{}^{-1}$$

and

$$\lambda_iC_i(\mathbf{r})/\phi(\mathbf{r})=\mu_i.$$

Equations (11), (10), and (5) are the system to be solved; it is clear that such a system is nonlinear. The system can be written in a more compact form as

$$\begin{aligned}
\frac{dx_1}{dt}&=A_{11}(t)x_1+A_{13}x_3\\
&\quad+A_{14}x_4\cdots+A_{1,m+2}x_{m+2}+E_1x_1x_2,\\
\frac{dx_2}{dt}&=A_{21}(t)x_1+A_{22}x_2+E_2x_1x_2,\\
\frac{dx_3}{dt}&=A_{31}(t)x_1+A_{33}x_3+E_3x_1x_2, \quad (12)\\
&\cdots\cdots\cdots\cdots\\
\frac{dx_{m+2}}{dt}&=A_{m+2,1}(t)x_1+A_{m+2,m+2}x_{m+2}+E_{m+2}x_1x_2,
\end{aligned}$$

or

$$\frac{d}{dt}\begin{Bmatrix}x_1\\x_2\\\cdot\\\cdot\\\cdot\\x_{m+2}\end{Bmatrix}=(A(t))\begin{Bmatrix}x_1\\x_2\\\cdot\\\cdot\\\cdot\\x_{m+2}\end{Bmatrix}+\begin{Bmatrix}E_1x_1x_2\\E_2x_1x_2\\\cdot\\\cdot\\\cdot\\E_{m+2}x_1x_2\end{Bmatrix}, \quad (13)$$

or

$$\frac{d}{dt}\mathbf{x}=A(t)\mathbf{x}+\mathbf{p}(\mathbf{x},t). \quad (14)$$

The nonlinear term is written as a function of $\mathbf{x}$ and t, even though it does not depend explicitly on t; the theorems which will be used can be applied to such general cases, however. In the equations above, symbols have the following meanings:

$$\begin{aligned}
&x_1=n(t),\quad x_2=T(t),\quad x_{2+k}=H_k(t),\\
&A_{11}(t)=k_{ex}{}^c(t),\\
&A_{22}=-\lambda,\\
&A_{1k}(k=3,4,\cdots,m+2)=\Lambda\mu_{k-2},\\
&A_{21}=\gamma,\\
&A_{j1}(j=3,4,\cdots,m+2)=\frac{k^c{}_{ex}(t)}{p}\beta_{j-2}\frac{\phi(r)}{C_{j-2}(r)}\sum_a, \quad (15)\\
&A_{kk}(k=3,\cdots,m+2)=-\lambda_{k-2},\\
&E_1=-\alpha,\quad E_2=0,\quad E_{2+k}=-\frac{\alpha}{p}\beta_k\frac{\phi(\mathbf{r})}{C_k(\mathbf{r})}\sum_a.
\end{aligned}$$

III. SOLUTION OF EQUATIONS

A. Preliminary Concepts

In order to make full use of existing theorems regarding differential equations (some of which have been obtained only within the last few years in the present form), it is necessary to put forth the concept of norm. The norm of an n dimensional vector z will be written as

(herein $n=m+2$)

$$\text{norm of } z=\|z(t)\|=\sum_{i=1}^{n}|z_i(t)|, \tag{16}$$

where terms on the extreme right represent the sum of the absolute values of the separate components. The norm of an $n\times n$ matrix A is written

$$\|A(t)\|=\sum_{j,k}^{n}|a_{jk}(t)|. \tag{17}$$

Existence theorems for systems of differential equations are usually stated in terms of norms. The theorems of interest here are those which provide a technique for the actual construction of the solution.

B. Iterative Technique

Equation (14) will be written as

$$\frac{d\mathbf{x}}{dt}=\mathbf{q}(\mathbf{x},t); \tag{18}$$

in order to demonstrate the conditions under which such an equation admits a unique solution, a method given by Lefshetz will be pointed out.[4] This method has a special advantage in that it provides a convenient method for the actual construction of the solution. Complete details of the proof are of no importance here. Suffice it to remark that instead of the conventional Lipshitz condition for a single differential equation, herein it is required that

$$\|\mathbf{q}(\mathbf{x},t)-\mathbf{q}(\mathbf{x}',t)\|\leq\epsilon\|x-x'\|, \quad \epsilon>0, \tag{19}$$

for every pair of relevant points $(\mathbf{x},t)$ and $(\mathbf{x}',t)$ in order that the technique given below be valid. The term relevant points means all values of $\mathbf{x}$ and t to be considered. When the inequality (Eq. 19) obtains, the solution obtained in the iterative fashion will converge uniformly to the solution of the original differential equation. The iterative technique provides various approximations as follows: One writes Eq. (18) as

$$\mathbf{x}=\int_{t_0}^{t}\mathbf{q}(\mathbf{x},\tau)d\tau. \tag{20}$$

The iterative scheme provides various degrees of approximation in that one starts with a so-called zeroth order approximation (value of $\mathbf{x}$ at time t_0) and substitutes this on the right hand side of Eq. (20) and performs the integration. This gives the first approximation, which is again substituted on the right hand side to get the second approximation; this procedure is then repeated until successive approximations agree within a previously prescribed amount. This iteration scheme may be represented as

$$\mathbf{x}^{(k+1)}=\mathbf{x}^{(0)}+\int_{t_0}^{t}\mathbf{q}(\mathbf{x}^{(k)},\tau)d\tau, \tag{21}$$

where the superscript represents the degree of approximation. This procedure is then continued for every relevant time interval.

It is fortunate that the existence (or nonexistence) of the condition given by Eq. (19) can be verified rather easily. One must examine each of the n components q_i of the vector $\mathbf{q}$ as follows:

$$q_i(\mathbf{x}',t)-q_i(\mathbf{x}'',t)=\sum_{j=1}^{n}\frac{\partial p_i(\mathbf{x},t)}{\partial x_j}(x_j'-x_j''), \tag{22}$$

where $\mathbf{x}'\leqslant\mathbf{x}\leqslant\mathbf{x}''$. If all of the $\partial p_i/\partial x_j$ are bounded, then one can take M as the upper bound. Multiply M by n, the number of components, which will yield the following inequality

$$\|\mathbf{q}(\mathbf{x}',t)-\mathbf{q}(\mathbf{x}'',t)\|\leq nM\|\mathbf{x}'-\mathbf{x}''\|, \tag{23}$$

and for $nM=\epsilon$, condition (19) is verified.

When computing machines are available which are able to follow a program such that a parameter is obtained and the procedure repeated with the new value of the parameter to get still a better value followed again by the same procedure with the most recent value, etc., it is clear that such a procedure as outlined above will be satisfactory for including the effect of delayed neutrons in reactor kinetics. In fact, in practice one will generally find iterative techniques very convenient (in spite of other methods).

C. Other Methods

A survey of other methods suitable for the numerical solution of such a problem may be found in a report by Bellman.[5] The basic idea in these methods consists of an examination of the conditions under which the solutions of the linear part of the equation are relevant to the complete equation; the solutions of the linear part of the equation are assumed to be known. At this point a few remarks regarding the solution of the linear equation are in order. Even though the method is well known, the specific application involved does not have extensive use.

The $n\times n$ matrix equation can be reduced to an nth order differential equation with time dependent coefficients. One then obtains n linearly independent solutions each of which may be considered as a vector in n-dimensional space. The solution can then be written as an $n\times n$ matrix each of whose columns is one of the n-dimensional linearly independent vectors. (It is clear that such a matrix has an inverse.) The linear part of

[4] S. Lefshetz, *Lectures on Differential Equations* (Princeton University Press, Princeton, New Jersey, 1948), pp. 23–30.

[5] R. Bellman, *A Survey of the Theory of the Boundedness, Stability, and Asymptotic Behavior of Solutions of Linear and Nonlinear Differential and Difference Equations* (Office of Naval Research report, January 1949); also, R. Bellman, *Stability Theory of Differential Equations* (McGraw-Hill Book Company, Inc., New York, 1953).

Eq. (14) or (18) may then be written as

$$\frac{d}{dt}Y=A(t)Y, \tag{24}$$

where Y is the $n\times n$ matrix referred to above. The solutions of Eqs. (18) and (24) are shown by Bellman to be linked through a Voterra-type integral equation

$$\mathbf{x}=\mathbf{y}+\int_0^t Y(t)Y^{-1}(\tau)\mathbf{p}(\mathbf{x},\tau)d\tau, \tag{25}$$

where $\mathbf{y}$ is a column in the solution of Eq. (24). If the $\mathbf{x}$ did not appear under the integral sign and one had only $\mathbf{p}(t)$, the integration could be performed directly (at least in principle) without resort to iteration. Bellman shows how to obtain upper bounds to the solutions for various restrictions on the functions. It is evident that the technique discussed under *B* is a more practical one.

IV. STABILITY OF SOLUTIONS

In a previous discussion,[2] it has been shown how the concept of reactor stability can be related to the stability of the solutions of the system of differential equations. Three forms of stability (according to Poincaré, Liapounoff, and Nyquist, respectively) were discussed. Each form of stability involved a perturbation of the initial conditions only. The conclusions drawn from the behavior of the solutions in the neighborhood of the origin prevailed for the stability of the system "in the large." No such general results can in general be inferred, for the system which includes delayed neutrons, from the study of the system which neglects the delayed neutrons. Even though neglecting the delayed neutrons gives an approximation to the actual system, the inferences drawn from an examination of the linear system may be invalidated because the effect of the delayed neutrons does not vanish in the neighborhood of the origin (i.e., at equilibrium). This point can be made clearer by rewriting Eq. (13) so that it shows explicitly one part which depends on the delayed neutrons and the other part which is independent of them. Equation (13) may be rewritten as

$$\frac{d}{dt}\begin{Bmatrix}x_1\\x_2\\x_3\\\cdot\\\cdot\\\cdot\\x_{m+2}\end{Bmatrix}=\begin{bmatrix}A_{11}(t) & 0 & 0 & \cdots & 0\\A_{21}(t) & A_{22} & 0 & \cdots & 0\\0 & 0 & 0 & \cdots & 0\\\cdot & \cdot & \cdot & \cdot & \cdot\\0 & 0 & 0 & \cdots & 0\end{bmatrix}\begin{Bmatrix}x_1\\x_2\\x_3\\\cdot\\\cdot\\x_{m+2}\end{Bmatrix}+\begin{Bmatrix}E_1x_1x_2\\E_2x_1x_2\\0\\\cdot\\\cdot\\0\end{Bmatrix}+\begin{bmatrix}0 & 0 & A_{13} & \cdots & A_{1,m+2}\\0 & 0 & 0 & \cdots & 0\\A_{31}(t) & 0 & A_{33} & \cdots & 0\\\cdot & \cdot & \cdot & \cdot & \cdot\\A_{m+2,1}(t) & 0 & 0 & \cdots & 0\end{bmatrix}\begin{Bmatrix}x_1\\x_2\\x_3\\\cdot\\\cdot\\x_{m+2}\end{Bmatrix}+\begin{Bmatrix}0\\0\\E_3x_1x_2\\\cdot\\\cdot\\E_{m+2}x_1x_2\end{Bmatrix}. \tag{26}$$

It is clear that Eq. (26) is identical to Eqs. (12) and (13). The first two terms on the right hand side of Eq. (26) are independent of the delayed neutrons whereas the latter two terms depend on them. Equation (26) may be written more compactly as

$$\frac{dx_i}{dt}=P_i(t,x_1,\cdots,x_{m+2})+D_i(t,x_1,\cdots x_{m+2}), \tag{27}$$

where the P_i depend only on prompt neutron behavior and D_i on delayed neutrons. For the purpose of stability discussions, one may consider the system as $dx_i/dt=P_i$; here the initial conditions are perturbed which affects the P_i and in addition the equations, themselves, are perturbed by a constantly acting disturbance D_i. The stability properties of $dx_i/dt=P_i$ can be learned by examining only the linear terms because in the neighborhood of the origin the nonlinear terms vanish more strongly. However, when constantly acting disturbances are present, these effects may override the linear terms even in the neighborhood of the origin. Thus, in general, the linear terms of the $dx_i/dt=P_i$ equation are not sufficient to determine the stability properties of the complete equation. In summary, one may say that even though the system in which the delayed neutrons are neglected in an approximation to the actual system, the stability properties of the approximate system may not be relevant for the actual system. In order to decide whether or not the system is stable, one must examine all of the terms in the P_i as well as the constantly acting disturbances D_i.

The concept of stability (which was stated above in a qualitative fashion) is given by Malkin as follows: The unperturbed state of motion (corresponding to the solution $x_1=x_2=\cdots=x_n=0$ of $dx_i/dt=P_i$) is defined as stable under constantly acting disturbances if for every positive number $\epsilon<H$, no matter how small, there exist two other positive numbers η_1 and η_2 (both depending on the magnitude of ϵ) such that all solutions of Eq. (27) with initial conditions (for $t=t_0$) satisfying the inequalities $|x_i^0|\leqslant\eta_1$ and for any D_i which satisfy the inequalities $|D_i(t,x_1,\cdots,x_n)|\leqslant\eta_2$ within the domain $\{t\geqslant t_0;\ |x_i|\leqslant\epsilon\}$, the relations $|x_i|<\epsilon$ hold for all $t\geqslant t_0$.

V. REMARKS

Techniques have been shown whereby the effect of the delayed neutrons on reactor kinetics may be taken into consideration. Conditions under which the equations may be integrated by iteration have been pointed out. The concept of stability has been formulated, and the

conditions are shown under which the stability properties of the systems with and without delayed neutrons can be related. The discussion given herein applies to systems of equations in which the time appears other than as the independent variable in the derivatives (nonautonomous systems), whereas previous discussions have been valid only for autonomous systems (i.e., time appears only as independent variable in derivatives).

In concluding, it is appropriate to point out the formal similarity that exists between the method of attack on this problem and on some approaches to time dependent perturbations in quantum mechanical problems. Equations of the form (20) and (21) are common;[6] in fact Eq. (21) is called various Born approximations. There are, however, two differences which should be mentioned. In the first place, the constants involved in the discussion herein are real, whereas complex constants (and functions) are admitted in quantum mechanical discussions. Secondly, the existence of the convergence of the "Born approximations" herein depends on the use of only finite dimensional vector spaces, whereas in quantum mechanical problems, spaces of an infinite number of dimensions are considered. Neither of these presents an essential distinction between the two problems. The fact that the norms are used in the essential proofs eliminates the apparent difficulty arising from the appearance of complex constants. All of the equations will have the same appearance. Reassuring statements regarding this specific point are given by Hurweicz.[7] Furthermore, an essential feature of a usable infinite dimensional space is that an arbitrary function can be expanded in a finite number of dimensions to a specified accuracy. Working with this number of dimensions preserves the treatment contained herein.

Hurwitz[8] has treated reactor kinetics from a different point of view.

[6] e.g., B. A. Lippmann and J. Schwinger, Phys. Rev. **79**, 469 (1950); E. C. Kemble, *Fundamental Principles of Quantum Mechanics* (McGraw-Hill Book Company, Inc., New York, 1937), pp. 427–431, 441.

[7] W. Hurweicz, *Ordinary Differential Equations in the Real Domain with Emphasis on Geometric Methods* (Brown University, Providence, Rhode Island, 1943), p. 64.

[8] H. Hurwitz, Jr., Nucleonics **5**, 61 (July, 1949).

JOURNAL OF APPLIED PHYSICS VOLUME 26, NUMBER 1 JANUARY, 1955

Qualitative Evaluation of Correlation Coefficients from Scatter Diagrams

THOMAS M. BURFORD
Department of Electrical Engineering, University of Wisconsin, Madison, Wisconsin

(Received April 19, 1954)

Middleton and Sugar have shown that the scatter diagram of two random electrical quantities may be oscilloscopically displayed, and that a simple relation exists between the resulting pattern and the correlation coefficient, if a Gaussian distribution is assumed. The present paper shows that qualitative information concerning the correlation coefficient may be found from the scatter diagram, no matter what distribution is assumed.

TWO random quantities, x and y, which are described simultaneously by their joint probability distribution, $f(x,y)$, have a correlation coefficient given by

$$\rho \equiv \frac{E[(x-m_1)(y-m_2)]}{[E(x-m_1)^2 E(y-m_2)^2]^{\frac{1}{2}}} \tag{1}$$

where*

$$E(x) \equiv m_1$$
$$E(y) \equiv m_2.$$

If x and y are given by a stochastic time process, then both $f(x,y)$ and therefore ρ would be functions of the separation of x and y in time. For the purposes of this discussion, however, $f(x,y)$ will be assumed fixed. Recent publications[1,2] have shown that ρ may be obtained quite easily from $f(x,y)$, if $f(x,y)$ is assumed to be Gaussian. It is the purpose of this paper to show that the result previously obtained for the special case of a Gaussian distribution is true for all physical distributions if a different interpretation is made of the quantities obtained from the scatter diagram.

The joint probability distribution $f(x,y)$ might be considered to be a surface in $[x,y,f(x,y)]$ space, but instead let us consider it as a mass density over the x,y plane.

Now, following Cramér,[3] we introduce a line, L, through the center of mass having the equation

$$(x-m_1)\sin\phi-(y-m_2)\cos\phi=0, \tag{2}$$

where ϕ is the angle between L and the positive x axis. The smallest distance between the mass containing point (x,y) and L is given by

$$d=(x-m_1)\sin\phi-(y-m_2)\cos\phi \tag{3}$$

* The operator E is defined as the expected value, that is, $E(H)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}Hf(x,y)dxdy$.

[1] Middleton, Quarterly Report No. 1, "Microwave noise study," Raytheon Manufacturing Company, May 1, 1953, to August 1, 1953.

[2] Sugar, J. Appl. Phys. **25**, 354 (1954).

[3] H. Cramér, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, New Jersey, 1951), p. 275.

and, therefore, the moment of inertia of the distribution about L is

$$E(d^2)=E[(x-m_1)\sin\phi-(y-m_2)\cos\phi]^2$$
$$=\sigma_1^2\sin^2\phi-2\rho\sigma_1\sigma_2\sin\phi\cos\phi+\sigma_2^2\cos^2\phi, \quad (4)$$

where

$$\sigma_1^2=E(x-m_1)^2$$
$$\sigma_2^2=E(x-m_2)^2.$$

In order to find the extremal values of $E(d^2)$, let

$$\frac{\partial}{\partial\phi}E(d^2)=0, \quad (5)$$

which gives the condition on ϕ as

$$\tan 2\phi=2\rho\sigma_1\sigma_2/\sigma_1^2-\sigma_2^2, \quad (6)$$

a result given also by Cramér.[4]

Evidently (6) defines two values of ϕ, $\pi/2$ radians apart, which determine the maximum and minimum values of the moment of inertia. Now, if a distance proportional to $[E(d^2)]^{-\frac{1}{2}}$ is marked off on L on each side of the mass center, a locus is determined in polar coordinates (r,ϕ):

$$r=\frac{C}{(\sigma_1^2\sin^2\phi-2\rho\sigma_1\sigma_2\sin\phi\cos\phi+\sigma_2^2\cos^2\phi)^{\frac{1}{2}}}. \quad (7)$$

If (7) is put into rectangular coordinates, the result is

$$\frac{x^2}{\sigma_1^2}-\frac{2\rho xy}{\sigma_1\sigma_2}+\frac{y^2}{\sigma_1^2}=\frac{C^2}{\sigma_1^2\sigma_2^2}. \quad (8)$$

Equation (8) evidently defines an ellipse, centered at the distribution's mass center, having the interesting property that a line joining two points on the ellipse through its center has a length inversely proportional to the square root of the moment of inertia of the probability distribution about that line. The ellipse considered here is directly analogous to Cauchy's momental ellipsoid as used in mechanics.[5]

The ratio of extremal values of inertia is

$$\frac{E(d^2)|_{\phi_1}}{E(d^2)|_{\phi_1+\pi/2}}=R, \quad (9)$$

where ϕ_1 is the angle defined by (6) that lies in the first quadrant. Substituting gives

$$R=\frac{\sigma_1^2\sin^2\phi-2\rho\sigma_1\sigma_2\sin\phi\cos\phi+\sigma_2^2\cos^2\phi}{\sigma_1^2\cos^2\phi+2\rho\sigma_1\sigma_2\sin\phi\cos\phi+\sigma_2^2\sin^2\phi} \quad (10)$$

and, making the convenient (though unnecessary) assumption that $\sigma_1=\sigma_2=\sigma$, which implies that $\phi_1=\pi/4$, and using (6), then

$$R=\frac{\sigma^2-\rho\sigma^2}{\sigma^2+\rho\sigma^2}=\frac{1-\rho}{1+\rho}. \quad (11)$$

Now, solving for ρ, we get

$$\rho=\frac{1-R}{1+R}; \quad (12)$$

and of necessity

$$|\rho|\leqq 1. \quad (13)$$

By using (9) we may write Eq. (12) as

$$\rho=\frac{E(d^2)|_{3\pi/4}-E(d^2)|_{\pi/4}}{2\sigma^2}. \quad (14)$$

Equation (12) above is similar in appearance to the expression previously found[1,2] for the Gaussian distribution. It should be recalled, however, that in (12) the quantity R is the ratio of the moment of inertia for $\phi=\pi/4$ to the moment of inertia for $\phi=3\pi/4$. The quantity occupying a similar place in the expression of Middleton and Sugar, on the other hand, is the square of the ratio of major to minor axes of an equi-probability ellipse. The similarity of the two results comes about as a result of the form of the Gaussian distribution in two dimensions, $g(x,y)$.

$$g(x,y)=\frac{1}{2\pi\sigma_1\sigma_2(1-\rho^2)}$$
$$\times\exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_1^2}-\frac{2\rho xy}{\sigma_1\sigma_2}+\frac{y^2}{\sigma_2^2}\right)\right]. \quad (15)$$

From the value of the exponent in (15) it is seen that a particular value of $g(x,y)$ implies a locus in the x,y plane that is an ellipse.

In conclusion, it has been shown that for any distribution the correlation coefficient may be found from the ratio of the extremal values of inertia. The inertial ratio will reduce, in the Gaussian case, to the square of the ratio of major to minor axes of any equi-probability ellipse. It is possible that the inertial ratio concept can be used to mechanize the process of finding correlations; but in any case, this ratio gives immediate qualitative information.

The author wishes to thank Professor V. C. Rideout of the University of Wisconsin for his advice in connection with this paper.

[4] H. Cramér, see reference 3.

[5] Whittaker, *Analytical Dynamics* (Dover Publications, New York, 1944), p. 124.

Closed Expansion of the Convolution Integral (A Generalization of Servomechanism Error Coefficients)

EDWARD ARTHURS AND LOUIS H. MARTIN
Servomechanisms Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts

(Received May 14, 1954)

This paper presents a closed expansion of the convolution integral, which is useful in approximating the response of a linear time invariant system to an arbitrary forcing function. In a sense it is a generalization of the well known error coefficient expansion frequently discussed in connection with servomechanisms. This expansion is helpful when the response of the system cannot be evaluated exactly (i.e., the input function is specified graphically or input function cannot be transformed conveniently). This expansion differs from the error coefficient expansion in that it is valid for an arbitrary forcing function and is in a closed form accounting for the entire response of the system, rather than in an open form of Taylor's series that either neglects or simply bounds a portion of the total response.

As a function of time, t, the response $r(t)$ of a stable linear system is related to the forcing function $f^0(t)$ by the convolution integral

$$r(t)=\int_{-\infty}^{t} f^0(x)w_0(t-x)dx, \tag{1}$$

where $w_0(t)$ is the response of the system to a unit impulse forcing function. The forcing function $f^0(t)$ is assumed to be continuous almost everywhere, suffering at most step discontinuities.[1] Thus,

$$f^0(t)=f_c{}^0(t)+\sum_{k=1}^{M_0}\Delta_{0k}u_{-1}(t-t_{0k}), \tag{2}$$

where $f_c{}^0(t)$ is the continuous portion of $f^0(t)$ and the summation represents the discontinuous portion. M_0 is the number of step discontinuities of $f^0(t)$, Δ_{0k} is the magnitude of the discontinuities, and t_{0k} the time of occurence of the kth discontinuity. In this paper $u_0(t)$ shall be used to represent a unit impulse, $u_{-1}(t)$ to represent a unit step, $u_{-2}(t)$ to represent a unit ramp, etc.[2]

By representing the forcing function $f^0(t)$ as a sum of continuous and discontinuous portions, the convolution integral can be expanded, as shown in Appendix I, to express the response of a stable linear system as the sum of three terms

$$r(t)=r_1(t)+r_2(t)+r_3(t). \tag{3}$$

These three terms are

$$r_1(t)=\sum_{j=0}^{N} c_j f^j(t), \tag{4}$$

$$r_2(t)=\sum_{j=0}^{N}\sum_{k=1}^{M_j}\Delta_{jk}w_{j+1}(t-t_{jk}) \tag{5}$$

and

$$r_3(t)=\int_{-\infty}^{t} f^{N+1}(x)w_{N+1}(t-x)dx, \tag{6}$$

where

$$f^j(t)=\frac{d}{dt}f_c{}^{j-1}(t), \tag{7}$$

$$w_{j+1}(t)=-u_{-1}(t)\int_t^{\infty} w_j(y)dy, \tag{8}$$

$$c_j=-w_{j+1}(0), \tag{9}$$

and $f^j(t)$ is the derivative of the continuous portion $f_c{}^{j-1}(t)$, of $f^{j-1}(t)$. Since $f_c{}^{j-1}(t)$ has no discontinuities, $f^j(t)$ can be expressed as the sum of a continuous portion and a summation of step discontinuities

$$f^j(t)=f_c{}^j(t)+\sum_{k=1}^{M_j}\Delta_{jk}u_{-1}(t-t_{jk}), \tag{10}$$

where M_j represents the number of discontinuities of $f^j(t)$, Δ_{0k} the magnitude of these discontinuities, and t_{jk} the time of occurrence of the kth discontinuity.

Equations (8) and (9) can also be expressed as

$$w_{j+1}(t)=-u_{-1}(t)\int_t^{\infty}\frac{(t-y)^j}{j!}w_0(y)dy, \tag{11}$$

$$c_j=\int_0^{\infty}\frac{(-y)^j}{j!}w_0(y)dy. \tag{12}$$

From Eq. (11) it is easy to see that $w_j(t)$ may be interpreted as the transient portion, $r_2(t)$, of the response $r(t)$ to $u_{-n}(t)$. From Eqs. (9) and (11) it is seen that c_j may be interpreted as the area under $w_j(t)$.

The usefulness of this expansion lies in the fact that in many problems $r_3(t)$ contributes a small portion of the total response $r(t)$. Consequently $r_3(t)$ may either be neglected or crudely approximated without introducing appreciable error in $r(t)$. In such cases, the process of convolution in Eq. (1) is replaced by differentiation and summation.

[1] If $f^0(t)$ contains impulses; their contribution can be added in separately.

[2] $u_{-n}(t)$ can be seen to be the inverse Laplace transform of $1/s^n$, where s is the complex frequency variable.

In this expansion $r_1(t)$ is equivalent to the familiar error coefficient expansion.[3] The second set of terms $r_2(t)$ accounts for possible discontinuities in the arbitrary forcing function and its derivatives. (Other discussions have assumed continuity of the derivatives of the forcing function employed in the error coefficient expansion, but Biernson[4] has pointed out that the effect of any discontinuities in a derivative can be accounted for by simply adding to the error coefficient expansion a corresponding transient term. These transient terms are essentially those accounted for by this second term.) The third term $r_3(t)$ is in the form of a convolution integral and clearly places in evidence the *exact error* involved in using a finite number of coefficients in the first term of this expansion.[5] Since this expansion is useful for many types of linear systems, and since it does not necessarily involve the expansion of an error function, we feel that it is more appropriate to call the coefficients in the first term of this expansion system coefficients rather than error coefficients.

To illustrate the application of Eq. (3), the azimuth angle response of a tracking system to a target moving on a crossover course will be computed. The line of sight input to the tracking system, $f^0(t)$, will be $f^0(t) = u_{-1}(t)\tan^{-1}(kt-1)$. The system impulse response, $w_0(t)$ will be taken as $w_0(t) = u_{-1}(t)(25/3)e^{-4t}\sin 3t$. Applying Eq. (7) we obtain

$$f^1(t) = u_{-1}(t)\frac{k}{1+(kt-1)^2},$$

$$f^2(t) = -u_{-1}(t)\frac{2k^2(kt-1)}{[1+(kt-1)^2]^2},$$

$$M_0=1, \quad \Delta_{01}=-\frac{\pi}{4}, \quad t_{01}=0,$$

$$M_1=1, \quad \Delta_{11}=\frac{k}{2}, \quad t_{11}=0.$$

Applying Eq. (39)[6] to $W_0(s)$ and inverse transforming we obtain

$$w_1(t) = -1.67u_{-1}(t)e^{-4t}\sin(3t+36.9°),$$

$$w_2(t) = 0.332u_{-1}(t)e^{-4t}\sin(3t+73.7°),$$

$$c_0=1, \quad c_1=-0.32.$$

For $N=1$ the three components of $r(t)$ are:

$$r_1(t) = u_{-1}(t)\left\{\tan^{-1}(kt-1) - \frac{0.32k}{1+(kt-1)^2}\right\}, \tag{13}$$

$$r_2(t) = u_{-1}(t)\{1.32e^{-4t}\sin(3t+36.9°) + 0.166ke^{-4t}\sin(3t+73.7°)\}, \tag{14}$$

$$r_3(t) = -u_{-1}(t)\int_0^t \frac{2k^2(kx-1)}{[1+(kx-1)^2]^2} 0.332e^{-4(t-x)} \times \sin(3[t-x]+73.7°)dx. \tag{15}$$

$r_3(t)$ can be evaluated by approximate methods and the error of approximation bounded, or $r_3(t)$ can be bounded. An upper bound for $r_3(t)$ is

$$|r_3(t)| < \int_0^t |f^3(x)|_{\max} 0.332e^{-4(t-x)}dx \tag{16}$$

evaluating the right hand side of the inequality

$$|r_3(t)| < 0.055k^2u_{-1}(t)[1-e^{-4t}]. \tag{17}$$

Hence if $k<1$, the total response is approximated very closely by $r_1(t)+r_2(t)$.

APPENDIX I. DERIVATION OF EXPANSION

In the following derivation it is assumed that $w_0(t)$ represents a stable system function, i.e.,

$$|w_0(t)| < A_0e^{-bt} \text{ almost everywhere,} \tag{18}$$

$$A_0, \quad b>0 \tag{19}$$

and

$$w_0(t)=0, \quad t<0. \tag{20}$$

The above conditions insure the existence of the w_j's and c_j's.

From Eqs. (8) and (18) it is obvious that

$$|w_j(t)| < A_je^{-bt} \text{ almost everywhere,} \tag{21}$$

$$A_j, \quad b>0 \tag{22}$$

since the integral of e^{-bt} is $-(1/b)e^{-bt}$.

The three terms of Eq. (3) can be derived from Eq. (1) by applying repeatedly the well known rule of integration by parts.

$$\int_a^b udv = uv\Big|_a^b - \int_a^b vdu. \tag{23}$$

Applying Eq. (23) to Eq. (1) and letting

$$u = f^0(x), \tag{24}$$

$$du = f^1(x)dx + \sum_{k=1}^{M_0}\Delta_{0k}u_0(x-t_{0k})dx, \tag{25}$$

$$dv = w_0(t-x)dx, \tag{26}$$

$$v = \int_{-\infty}^{x} w_0(t-y)dy \tag{27}$$

[3] See for example James, Nichols, and Philips, *Theory of Servomechanisms* (McGraw-Hill Book Company, Inc., New York, 1947), p. 147.

[4] G. A. Biernson, "A simple method for calculating the time response of a system to an arbitrary input," Engineering Report 7138-R-3, Servomechanisms Laboratory, M.I.T., Cambridge, Massachusetts, 1954.

[5] An error coefficient expansion presented by John L. Bower, J. Appl. Phys. **21**, 723 July, 1950, simply placed a bound on the error incurred by using a finite number of coefficients. The exact error accounted for by $r_2(t)$ is the *lowest* bound that can be found for the error incurred by using a finite number of coefficients.

[6] See Appendix II.

we find that

$$r(t)=f^0(t)\int_{-\infty}^{t} w_0(t-y)dy-\sum_{k=1}^{M_0}\Delta_{0k}\int_{-\infty}^{t_{0k}} w_0(t-y)dy$$

$$-\int_{-\infty}^{t} f^1(x)dx\int_{-\infty}^{x} w_0(t-y)dy. \quad (28)$$

Equation (28) may be simplified in terms of Eqs. (8) and (9) obtaining

$$r(t)=c_0f^0(t)+\sum_{k=1}^{M_0}\Delta_{0k}w_1(t-t_{0k})$$

$$+\int_{-\infty}^{t} f^1(x)w_1(t-x)dx. \quad (29)$$

The last term of Eq. (29) is in the form of a convolution integral and it can be expanded in the same manner as Eq. (1) has been to obtain Eq. (29). If this process is continued, the three-term expansion of Eq. (3) will result.

The relationship indicated in Eq. (11) may be derived as follows: Let

$$-\int_t^\infty \frac{(t-y)^j}{j!}w_0(y)dy=-\int_t^\infty udv, \quad (30)$$

$$u=\frac{(t-y)^j}{j!}, \quad (31)$$

and

$$dv=w_0(y)dy. \quad (32)$$

Applying Eq. (17) it follows that[7]

$$uv\Big|_t^\infty=\left[\frac{(t-y)^j}{j!}w_0(y)\right]_t^\infty.$$

$$-\int_t^\infty \frac{(t-y)^j}{j!}w_0(y)dy=-\int_t^\infty \frac{(t-y)^{j-1}}{(j-1)!}w_1(y)dy. \quad (33)$$

[7] Since $w_0(y)\to 0$ exponentially as $y\to\infty$, the contribution of the upper limit is zero.

Repeating this process m times we obtain

$$-\int_t^\infty \frac{(t-y)^j}{j!}w_0(y)dy=-\int_t^\infty \frac{(t-y)^{j-m}}{(j-m)!}w_m(y)dy. \quad (34)$$

Then, for $j=m$,

$$-\int_t^\infty \frac{(t-y)^j}{j!}w_0(y)dy=-\int_t^\infty w_j(y)dy=w_{j+1}(t). \quad (35)$$

APPENDIX II. SOME USEFUL RELATIONSHIPS IN THE FREQUENCY DOMAIN

It is sometimes useful to consider this closed expansion in the frequency domain as well as the time domain. If we let $W_j(s)$ be the Laplace transform of $w_j(t)$, by definition

$$W_j(s)=\int_0^\infty w_j(t)e^{-st}dt. \quad (36)$$

Integrating by parts we obtain

$$W_j(s)=-\left[e^{-st}\int_t^\infty w_j(y)dy\right]_0^\infty$$

$$-s\int_0^\infty e^{-st}dt\int_t^\infty w_j(y)dy \quad (37)$$

or

$$W_j(s)=c_j+sW_{j+1}(s). \quad (38)$$

Starting with $j=0$ and repeatedly applying Eq. (38), it follows that

$$W_0(s)=c_0+c_1s+\cdots+c_ns^n+s^{n+1}W_{n+1}(s). \quad (39)$$

If $W_0(s)$ is a rational algebraic function of s,

$$W_0(s)=\frac{a_0+a_1s+\cdots+a_qs^q}{b_0+b_1s+\cdots b_ps^p}=\frac{Q(s)}{P(s)}. \quad (40)$$

It is obvious from Eq. (39) that the c_j and $W_j(s)$ can be computed by the simple process of dividing $P(s)$ into $Q(s)$.

JOURNAL OF APPLIED PHYSICS VOLUME 26, NUMBER 1 JANUARY, 1955

Measurements of Electrical Polarization in Thin Dielectric Materials*

R. W. TYLER, J. H. WEBB, AND W. C. YORK
Research Laboratories, Eastman Kodak Company, Rochester, New York
(Received December 7, 1953)

Electrical polarization in thin, dielectric sheet material induced by handling or other means is not readily observed because the external field from a dipole charge layer is very weak compared with the field from net charges that are usually present. A method is described for measuring polarization in the presence of net charges by placing the dielectric sheet in contact with a grounded metal backing plate and by placing a field meter at a short distance in front of the dielectric. The grounded metal backing plate has the effect of almost completely annulling the external field of any net charge by induction, and of doubling the external field of a dipole charge layer of polarization.

INTRODUCTION

THE purpose of this paper is to describe a method of measuring polarization in thin, dielectric sheet material when there is a significant amount of net charge of one sign also present on the sheet. The method depends essentially upon the measurement of the potential of the outer surface of a polarized dielectric sheet when backed by a grounded metal plate. The backing plate greatly reduces the external field of any net charge and doubles the external field of polarization.

The type of dielectric material with which this paper is chiefly concerned is cellulose acetate film support. Operations in a film manufacturing plant involve the handling of this material in both sheet form and in continuous-strip form where, in the latter case, the material moves over systems of support rollers. It is handled in widths which may vary from 16 millimeters to 40 inches. In handling operations, the film support often becomes highly charged with static electricity and, at times, polarized. If polarized film is wound into a large roll, or stacked in piles with many layers superposed, high electric fields result from the additive effects of the polarized layers.

A word should be included here concerning the nature of electrical polarization occurring in dielectric sheet material and how this property can be specified. When an electric field is applied to a dielectric, the positive and negative charges of its molecular units are displaced in opposite directions. Part of this shift takes place instantaneously on applying the field and is relieved as quickly on removing the field. However, there is another part of this shift which is absorptive and takes place slowly when the field is applied and dies out slowly after the field is removed. Residual effects in the dielectric of this nature constitute a type of polarization subject to measurement by the method outlined in this paper. Polarization of this type can be specified quantitatively as the dielectric moment per unit volume of the dielectric. Owing to the additivity of distributed polarization in a dielectric, a volume distribution of polarization in a dielectric sheet can always be duplicated in its external effects by equal and opposite charge layers on opposing surfaces of the dielectric sheet, provided the product of the charge density and the thickness of the dielectric has the appropriate electric moment per unit volume. Accordingly, it is often convenient to specify polarization in terms of an equivalent net charge density, $\pm\sigma_n$, on the two surfaces. In the present paper, polarization will be considered to arise from equal and opposite charge densities deliberately placed on the two surfaces of the dielectric sheet. If free charge densities, σ and $-\sigma$, are placed directly on the two opposing surfaces, a field is established across the dielectric. The field inside the dielectric will be $4\pi\sigma/k$, where k is the dielectric constant of the material. If there were no dielectric material between the charge layers, the field would have the value $4\pi\sigma$. Thus, the field inside the dielectric, owing to reactive effects, is reduced by an amount $4\pi\sigma(1-1/k)$, which is equivalent to the field produced by concealed charge densities $\pm\sigma(1-1/k)$. Therefore, when free charge densities σ and $-\sigma$ are placed on the surfaces of a dielectric sheet, the net effective charge density σ_n to be used in calculating the electrical field, whether inside or outside the dielectric, is σ/k.

It is to be emphasized that the method described here for measuring polarization will respond equally well to a volume distribution of dipoles throughout the dielectric or to dipole charge layers confined to the surfaces of the dielectric. In all cases studied here, the polarization is produced and specified in terms of dipole surface charge layers, for simplicity.

This paper is divided into two parts. The first part deals with the theoretical aspects of fields emanating from charged and polarized dielectric sheets and with the theory of the proposed method of measuring polarization. The second part of the paper gives the results of experimental measurements of polarization carried out by the new method and examples to illustrate its usefulness.

THEORETICAL PART

The electric field intensity at a point P on the axis of a circular plate uniformly charged with electricity of one sign is

$$E=2\pi\sigma\left[1-\frac{x/a}{(1+(x/a)^2)^{\frac{1}{2}}}\right]\frac{\text{es volts}}{\text{cm}}, \qquad (1)$$

* Communication No. 1629 from the Kodak Research Laboratories.

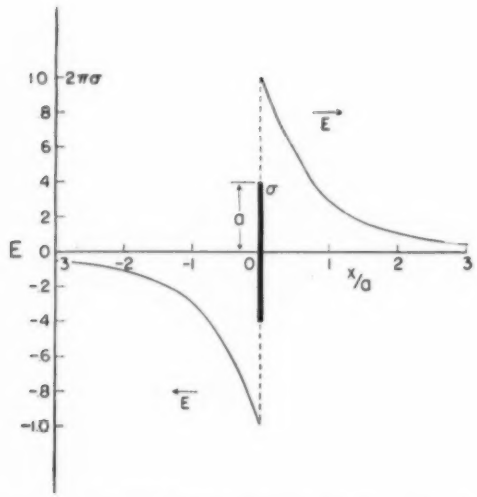


FIG. 1. Electrical field intensity about plane, circular, layer of charge of density σ, and radius a.

where σ is the charge per unit area of the plate in electrostatic units, x is the distance in centimeters of the point P from the center of the plate, and a is the radius of the plate in centimeters. The intensity can be converted to volts per centimeter by multiplying by the factor 300, which is the relation between es units and practical units. The intensities have the same absolute magnitudes but opposite signs on the two sides of the charged plate, as shown in Fig. 1.

The density of charge σ that can be placed on a single plate in air is limited by the breakdown strength of air, which is of the order of 30 000 volts per cm. In practice, this figure cannot be realized because of corona leakage. However, using this maximum figure, the limiting amount of charge σ that can be placed on such a plate can be calculated from Eq. (1), with x/a set equal to zero, and is found to be 15.9 esu/cm^2.

The field intensity about a dipole charge layer made up of equal and opposite charge densities on the two faces of a dielectric sheet can be obtained by addition of two expressions of the type (1) for two circular layers of charge of opposite sign. If charge densities σ and $-\sigma$ esu/cm^2 are placed on the two opposing faces of a dielectric sheet of thickness t cm and radius a cm, and the zero point of potential is chosen as the midpoint between the two surfaces, the field intensity at internal and external points of the sheet can be readily calculated. The field on the axis within the dielectric will be

$$E_i=-\frac{2\pi\sigma}{k}\left[2-\frac{x}{(x^2+a^2)^{\frac{1}{2}}}-\frac{t-x}{[(x-t)^2+a^2]^{\frac{1}{2}}}\right]. \quad (2)$$

At either surface, i.e., for $x=0$ or $x=t$, this field becomes

$$E_i=-\frac{2\pi\sigma}{k}\left[2-\frac{t}{(t^2+a^2)^{\frac{1}{2}}}\right], \quad (3)$$

and for the condition $t/a\ll 1$, the internal field reduces to the expression

$$E_i=-\frac{4\pi\sigma}{k}. \quad (4)$$

At points to the right or left of the double charge layer, the fields have the same sign and magnitude but have opposite sign to the internal field. The external field of the double charge layer is given by the expression†

$$E_e=\frac{2\pi\sigma}{k}\left[\frac{x}{(x^2+a^2)^{\frac{1}{2}}}-\frac{x-t}{[(x-t)^2+a^2]^{\frac{1}{2}}}\right]. \quad (5)$$

Equation (5) reduces to the approximate form

$$E_e=\frac{2\pi\sigma}{k}\frac{t}{a}\left[\frac{1-(x/a)^2/1+(x/a)^2}{(1+(x/a)^2)^{\frac{1}{2}}}\right] \quad (6)$$

for small values of t/a such that terms containing $(t/a)^2$ can be neglected. A curve of field intensity drawn according to Eq. (6) is shown in Fig. 2. The maximum value of intensity which occurs at either surface of the dipole charge layer, for $x=0$ or $x=t$, is approximately

$$E_e=\frac{2\pi\sigma}{k}\frac{t}{a}, \quad (7)$$

that is, reduced by the factor t/a relative to that for a single layer of charge of the same density σ. Thus, if the charge density of 15.9 esu/cm^2, which gives breakdown intensity in air for the single charge layer, is placed on each of the two sides of the dielectric sheet, the maxi-

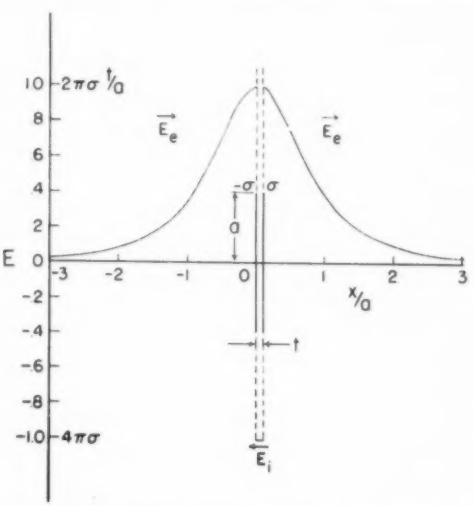


FIG. 2. Electrical field intensity about plane dipole charge layer of densities σ and $-\sigma$ separated by distance t cm.

† It may appear incorrect to use the dielectric constant k in the expression for the external field of the dipole charge layer. That the charge $\pm\sigma/k$ is the correct value to use for the external field, as well as for the internal field, is readily established by the requirement of continuity of potential at the surface of the dielectric.

mum external field from the dipole charge layer will be

$$\frac{2\pi\sigma}{k}\frac{t}{a}(300)=48.6 \text{ volts/cm},$$

in which the constants t and k have been given the respective values 0.013 cm and 4 for acetate film support and the radius a has been set equal to 2 cm. It is seen that the external field intensity from the dipole charge layer is extremely small in comparison with that from a net charge of the same density. With a dipole charge layer on the dielectric sheet, the main field intensity is internal, and the limiting charge density is set by the breakdown strength of the dielectric rather than by the breakdown strength of the surrounding air. For acetate film support of 0.013-cm thickness, the breakdown potential is approximately 11 100 volts. The charge density $\pm\sigma$ on the two surfaces required to give this potential is readily calculated, using the relation (4) multiplied by t,

$$\frac{4\pi\sigma}{k}t(300)=11\,100 \text{ volts},$$

and is found to be 907 esu/cm². By using this value of σ in Eq. (7), the external field of the dipole layer is found to be 2770 volts/cm at the surface of the dielectric. At increasing distances x from the surface, this field will diminish in accordance with Eq. (6), which is illustrated in Fig. 2. Also, the external field of a dipole charge layer decreases as its area is increased, the field approaching zero as the radius a increases to large values relative to t. For these reasons, and also because there is usually a sufficient amount of net charge of one sign present to mask the external field of a dipole layer completely, it is not practical to measure polarization of a dielectric sheet by observing the external field of the isolated layer. A practical method for measuring the polarization of a dielectric sheet material of any size for the case in which both net charge and polarization are present, will now be described.

First, let us consider the field from net charge alone on the dielectric sheet. We have already seen that the maximum unbalanced net charge that can exist on a dielectric sheet in air is 15.9 esu/cm². If a dielectric sheet with this net charge density is placed in contact with a grounded metal backing plate, an equal and opposite charge density layer will be induced in the metal. If the net charge layer on the dielectric is assumed to reside in the extreme position of the outer face of the dielectric of thickness t cm, then the total distance between the net charge and its image in the metal will be $2t$ cm. Under these conditions, the external field intensity from the dipole charge layer at the front surface of the dielectric sheet will be

$$E=\frac{2\pi\sigma}{k}\left(\frac{2t}{a}\right)300\frac{\text{volts}}{\text{cm}}. \qquad (8)$$

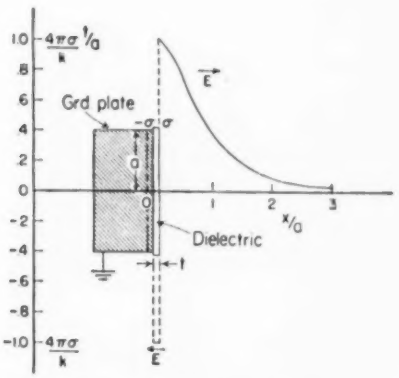


FIG. 3. External electric field intensity of charged dielectric sheet in contact with grounded metal plate.

For a net charge density of 15.9 esu/cm², and with $k=4$, $t=0.013$ cm, and $a=2$ cm, the axial field intensity at the outer surface of the dielectric will be 97.2 volts/cm. For a smaller net charge and/or a larger surface area of the dielectric sheet, the field will be correspondingly less. Also, the field will decrease with distance x from the dielectric surface in accordance with Eq. (6). With the usual amount of net charge encountered in practice, the external field from net charge on the backed dielectric sheet will be much smaller than the value of 97.2 volts/cm.

Second, let us consider the effect of surface dipole charges on such a sheet in contact with metal as shown in Fig. 3. The field from the inner surface is cancelled by that of the charge which it induces on the metal. The external field from the outer charge layer and its image will be given by Eq. (8) as before. We have seen that the maximum value of σ which this dielectric will support is 907 esu/cm², if $t=0.013$ and $k=4$. For a radius $a=2$ cm, this gives a field intensity of 5540 volts/cm.

It is evident from the foregoing analysis and calculations that a grounded metal backing plate greatly reduces the field of net charge on a dielectric sheet and at the same time doubles the external field of polarization that is present. Since ordinary dielectric sheet materials can support dipole charge densities one or two orders of magnitude greater than net charge, it may be seen that the grounded backing plate introduces a marked advantage for measuring a field of polarization in the presence of net charge. This is emphasized in Table I, which summarizes the values of field intensity calculated for net charge and dipole charge layers with and without the backing plate.

Notwithstanding the relatively great enhancement of the field of dipole charge layers by the metal backing plate, there remains one serious obstacle in the way of precise measurement of dipole fields as outlined so far. The external dipole field depends upon the area of the dielectric sheet, as may be seen from the appearance of the radius a in the expression (8). It will now be shown that this factor is automatically eliminated in the pres-

TABLE I. Field intensity for varying net charges and dipole charge layers.

	Field from max. net charge $\sigma=15.9$ esu/cm² (volts/cm)	Field from dipole charge $\pm\sigma=15.9$ esu/cm² (volts/cm)	Field from max. dipole charge $\pm\sigma=907$ esu/cm² (volts/cm)
Isolated dielectric sheet	30 000	48.6	2770
Dielectric in contact with grounded metal plate	0–97.2	97.2	5540

ent method of field measurement by using a fieldmeter placed at a close distance in front of the dielectric sheet.

It is assumed again that the dielectric sheet with charge densities σ and $-\sigma$ esu/cm² on its two faces is placed in contact with the grounded metal plate. In addition, we consider a second grounded metal plate, which represents the flat face of the fieldmeter and a surrounding metal shield, placed at a distance s from the front surface of the dielectric sheet and parallel to it, as shown in Fig. 4. As drawn here, the distances t and s are greatly exaggerated relative to the diameter of the dielectric sheet, actual values of t and s being of the order of 0.013 cm and 0.477 cm, respectively, while a is a few cm. From the standpoint of the outer surface of the dielectric sheet with charge σ, the grounded metal backing surface and the grounded metal fieldmeter surface form the outer plates of two condensers in parallel, the charge layer σ esu/cm² acting as the intermediate common plate of the two condensers. The condenser across the dielectric sheet of thickness t and dielectric constant k will have a capacity

$$C_t=\frac{kA}{4\pi t}, \tag{9}$$

while the condenser across the air space s will have a capacity

$$C_s=\frac{A}{4\pi s} \tag{10}$$

in which A represents the common area of the two condensers. The distances t and s are sufficiently small so that the fields across the condensers will be uniform.

The measurement of the polarization of the dielectric sheet requires essentially that the magnitude of the dipole charge σ on the outer surface of the dielectric sheet be determined when there is no fieldmeter present. Since the charge density σ is related to the potential of the outer surface of the dielectric sheet by the relation,

$$V_1=\frac{Q}{C_t}=\frac{4\pi\sigma}{k}t \text{ es volts}, \tag{11}$$

it will be sufficient to determine the potential V_1 when there is no fieldmeter present. As it is impossible to measure the surface potential of the dielectric sheet without the fieldmeter, it must be shown that the presence of the fieldmeter has negligible influence on the potential V_1 of the dielectric surface. It can readily be shown that the ratio of the voltages with and without the fieldmeter present is

$$\frac{V_1}{V_2}=\frac{C_t+C_s}{C_t}=1+\frac{t}{sk}. \tag{12}$$

If the constants t, s, and k have the values 0.013 cm, 0.477 cm, and 4, respectively, then

$$\frac{V_1}{V_2}=1.0068.$$

The presence of the fieldmeter in this case has changed the potential of the dielectric surface by a negligible amount. We shall now determine the magnitude of the field from the dielectric surface measured by the fieldmeter. The field intensity lines from the net charge layer σ of the dielectric surface terminate in part on an induced charge layer of density σ_t in the metal backing plate, and in part on an induced charge layer of density σ_s in the fieldmeter face, where the sum of σ_s and σ_t is equal to σ. The relative magnitudes of the charge densities σ_t and σ_s can be readily determined to be

$$\sigma_t=\frac{\sigma}{1+t/sk}, \quad \sigma_s=\sigma\frac{t/sk}{1+t/sk}. \tag{13}$$

With the constants $t=0.013$ cm, $s=0.477$ cm, and $k=4$, the quantity t/sk is equal to 0.0068, and $\sigma_t=0.9932\sigma$ and $\sigma_s=0.0068\sigma$. By using the value for σ of 327 esu/cm² corresponding to a potential of 4000 volts for the outer face of the dielectric (without fieldmeter present), the uniform field intensity in the space before the fieldmeter is

$$E_s=4\pi\sigma_s=8220 \text{ volts/cm}. \tag{14}$$

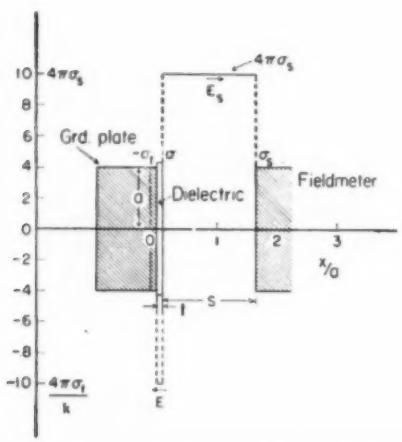


FIG. 4. Electric field intensity of charged dielectric sheet in contact with grounded metal plate with fieldmeter present.

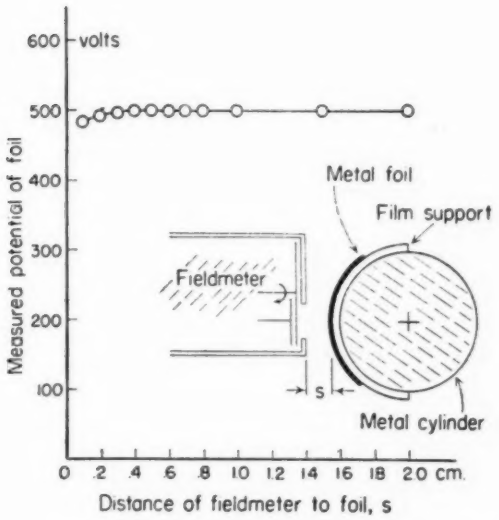


FIG. 5. Curve showing effect on potential of charged, insulated metal foil due to proximity of fieldmeter.

This field is shown plotted in Fig. 4 as a straight horizontal line. It should be noted that this field is independent of the size of the dielectric sheet, and depends only upon the potential of its outer surface. The potential of the outer surface of the dielectric sheet,when the fieldmeter is present, can be calculated from the voltage drop across the dielectric sheet of thickness t, or across the fieldmeter separation gap s. Assuming a value for σ of 327 esu/cm², corresponding to a potential of the dielectric surface of 4000 volts without the fieldmeter, we find

$$V_t=\frac{4\pi\sigma_t t}{k}300=4\pi\sigma_a s(300)=3975 \text{ volts}, \qquad (15)$$

again showing the very slight effect of the presence of the fieldmeter on the potential of the dielectric surface. The potential of the outer surface of the dielectric sheet will be affected only when the fieldmeter is brought extremely close to the dielectric. This has been confirmed by direct experiment, as shown in Fig. 5. A sheet of metal foil was placed on one surface of a piece of the acetate film support of thickness 0.013 cm and dielectric constant k. The opposite face of the dielectric support was placed in contact with the grounded metal plate. The fieldmeter was mounted in front of the dielectric support so that the distance between the foil and fieldmeter could be varied. The foil was raised to a potential of 500 volts and then disconnected from the source of supply voltage. Fieldmeter readings of the metal foil were made at distances over the range, 0.1 cm to 2 cm. Comparison of these readings with those for the foil connected to the 500-volt supply, show the effect of the presence of the fieldmeter on the potential of the insulated foil. It may be seen that the voltage dropped slightly only for distances below 0.4 cm. All measurements in the experimental work to follow were made with the fieldmeter distances of 0.477 cm.

It may be seen that the method of measuring polarization outlined here depends upon measuring the potential of the outer surface of a dielectric sheet backed by a grounded metal plate. Thus, the fieldmeter can be directly calibrated by setting it at a fixed distance in front of a metal foil held at a known potential, a method which rules out various ambiguities and eliminates the influence of geometrical factors.

EXPERIMENTAL PART

Apparatus and Method

All measurements of electric field intensity in the present study were made by means of a generating electrostatic voltmeter, which is referred to as a fieldmeter.[1] A schematic drawing of this instrument is shown in Fig. 6. Its essential elements consist of two sector disks having the same number of blades and cut to the same geometrical pattern. One of the sector disks remains stationary and is connected to ground through a resistance R of the order of 10^6 ohms. The other sector, which rotates, is placed closely adjacent to the stationary sector disk and is electrically grounded, or biased slightly above ground potential, for reasons to be given below. As the shielding sector rotates, an alternating current is set up in the circuit between the stationary sector and the ground. The resulting alternating voltage across the resistance R is measured by a vacuum-tube voltmeter of high sensitivity.

An apparatus for continuously recording the electrical polarization of long strips of 35 mm photographic film support is shown in Fig. 7. The film is passed over a metal roller 4 inches in diameter and 4 inches long, mounted in ball-bearing supports. The fieldmeter is placed in close proximity to the film support passing

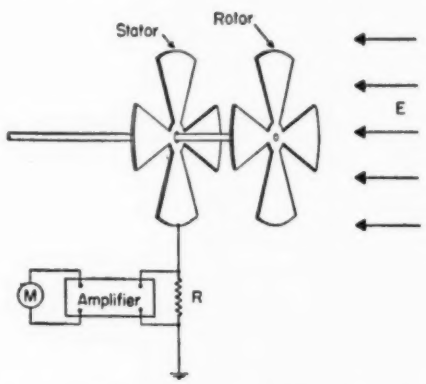


FIG. 6. Schematic drawing of fieldmeter used for measuring film-surface potentials.

[1] A. Matthias, Electrizatätswirtschaft **25**, 297 (1926); Ross Gunn, Phys. Rev. **40**, 307 (1932); G. P. Harnwell and S. N. VanVoorhis, Rev. Sci. Instr. **4**, 540 (1933); R. Pinoir, Compt. rend. **205**, 1369 (1937); H. W. Cleveland, J. Soc. Motion Picture Television Engrs. **55**, 37 (1950).

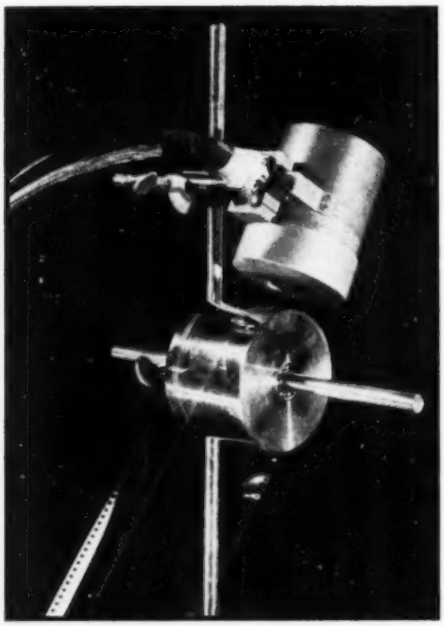


FIG. 7. Polarization detector consisting of fieldmeter in proximity to grounded metal roller.

over this roller, as shown in Fig. 7. In measuring polarization in this way, it is necessary to shield the aperture of the fieldmeter from the strong electrostatic fields arising from net charges on the film as it approaches and leaves the roller. This is accomplished by supplying the fieldmeter with a cap having a $\frac{3}{4}$-inch diameter aperture placed directly over one of the blades of the stationary sector disk and positioning this aperture at a very small distance ($\frac{3}{16}$ inch) from the roller at the center position of the 90-degree wrap of the film on the roller. The polarization detector was calibrated by determining the reading of the fieldmeter when a given potential was applied to a conducting metal foil in the position of the exposed surface of the film. By adjustment of the amplifier, the fieldmeter could be made to read directly the voltage of a surface at a fixed separation. The amplifier used gave a linear response and was capable of operating a recording 5 milliampere meter. The fieldmeter was biased to a center zero by applying a dc potential to its rotating sector disk, which permitted the measurement of either negative or positive potentials of the film surface.

The polarization detector was mounted as an integral part of a film-winding apparatus in which the film support could be unwound from a supply roll, passed over a series of rollers, including the detector roller, and finally wound up on a take-up roll. In addition to the polarization detector, a second fieldmeter was mounted so that the potential arising from net charge could be measured for the strip of film in free space as it passed from the supply roll to the polarization detector.

Results of Performance Tests

The dielectric material used in all of the following tests was cellulose acetate film support without emulsion coating.

A short strip of film support was given a net charge by passing it through the corona discharge between a grounded metal plate and a needle maintained at +5000 volts direct current. When this film was held before the fieldmeter, the potential of its surface was found to be +1200 volts, regardless of which face of the film was presented to the fieldmeter. When the filmstrip was held firmly in contact with the roller of the polarization detector, the meter reading was +2 volts from one face and +1 volt from the other, demonstrating the low response of the polarization detector to net charge on the film.

A second experiment was carried out using a piece of film which, in effect, had been polarized by placing charges of opposite signs on its two respective surfaces.

The strip of film was charged by laying it on a small, grounded block of brass, and spraying the charge onto its upper surface with a high-potential needle. When the film was stripped from the block, a discharge occurred which placed a neutralizing charge on the opposite face of the film to form the dipole. As the discharge was never complete, there always remained a net charge, which had the same sign as the needle potential used. By using different needle potentials, and various needle-film distances, a variation in dipole strength was obtained.

The potential of the film in space was measured with an electrometer and a potential probe. The potential of the outer face of the film, when placed on a grounded roller, was measured in the polarization detector. This was done for both surfaces. The results of this experiment are shown in Table II.

The dipole charge densities appearing in this table were computed from Eq. (11), with the values of $t=0.013$ cm and $k=4$. The opposite signs of the potential from the two sides demonstrated that the film carried an effective polarization, owing to opposite charge densities on its two surfaces. The approximate equality of the measurements on the two sides shows again that the detector is not materially influenced by the relatively large net charges present.

The polarization detector was tested next with

TABLE II. Potentials of net charge in space and dipole surface potentials in polarization detector (strips of cellulose acetate film with double charge layers).

Sample No.	Potential in space, volts	Dipole potentials, volts		Density of dipole charges, esu/cm²	
1	+19 000	+650	−600	+53	−49
2	+20 000	+1900	−2300	+150	−190
3	+22 000	+2400	−2200	+190	−180
4	+21 000	+1300	−1150	+100	−93
5	+23 000	+950	−900	+77	−73

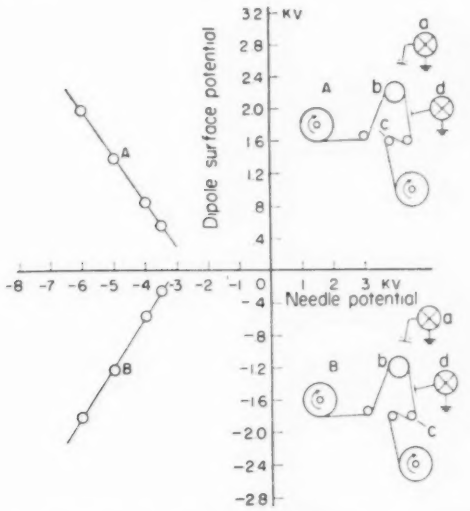


FIG. 8. Response of polarization detector to moving film subjected to needle discharge at grounded roller. (a) Fieldmeter; (b) Grounded metal roller; (c) Charged needle; (d) Electrometer and probe.

moving 35-mm film polarized continuously as it passed through the film-winding apparatus. After leaving the supply stock roll, the film was passed over two grounded metal rollers in succession, as shown in Fig. 8. If a needle point, charged to a high negative potential, is placed opposite one of these rollers, a negative charge is sprayed on the film. The charge density acquired by the outer surface of the film under these conditions can be very large, because of the high capacity between the film and the grounded roller surface. As the film leaves the roller, this capacity is greatly reduced, and the resulting high electric field between the film and roller produces a discharge to the inner face of the film. This charge on the inner surface of the film is opposite in sign to that sprayed on the outer surface, and therefore the film acquires a dipole charge layer. As the contacting surface of the film is reversed on the two successive rollers, the direction of the effective polarization of the film can be reversed by placing the needle opposite one or the other of the two rollers. Curves showing variation of effective polarization as a function of needle potential are presented in Fig. 8. The two curves correspond to application of negative charge to opposite surfaces of the film. The needle potential was varied in each case from -3.5 kv to -6 kv. The resulting potential measurements of the outer surface of the film are shown by the ordinate values, which vary from a few hundred volts to about 2000 volts.

The next experiment consisted in putting a net charge on a continuously moving strip of film by placing the discharge needle opposite the filmstrip without the backing roller. A fieldmeter was placed so as to read the potential of the moving filmstrip in free space before it reached the polarization detector. This potential, which arises from net charge on the film, is plotted as a function of the charging needle potential in Fig. 9A. The needle potential was varied from -9 kv, to $+9$ kv, and the resulting potential of the charged strip of film was observed to change from about -8 kv to $+7$ kv. This curve illustrates the high potential of a strip of film having relatively small amounts of net charge.

In Fig. 9B are shown the readings obtained on the polarization detector when the film had been given the net charge described above. This curve shows a polarization potential of about -245 volts over the entire range of needle potentials employed. The small slope of this curve indicates the slight effect of the net charge on the polarization readings.

The foregoing experiments confirm earlier statements that, in general, polarization or dipole layer charges can be much larger than net charge, and that the fields from these two types of electrification in a dielectric sheet can be readily distinguished by separate measurements of the electric fields from the film under the two conditions: (1) when in free space, and (2) when in contact with a grounded metal plate.

The results shown in Fig. 10 illustrate one practical application for the polarization detector in testing the electrostatic condition of photographic film support. A 1000-ft roll of photographic film support was rewound on a film-winding machine at a constant speed of 94 feet per minute under tension of 1000 grams. A record of the strip potential arising from the net charge on the film is shown in the upper chart. The lower chart shows a record of the polarization potential, as measured when the film was backed by the grounded metal roller. Whereas the strip potential is erratic and reaches very high positive and negative values in certain regions (4 to 5 kv), the polarization potential is of smaller magnitude

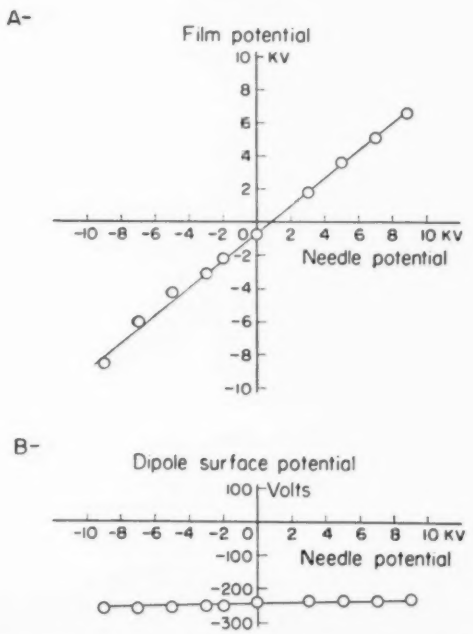


FIG. 9. (A) Potential of film and (B) polarization of film subjected to needle charge in free space.

and much greater uniformity. This record of polarization is typical of the polarization frequently encountered in practice, when film is wound and rewound under practical handling conditions. However, if film with this degree of polarization is wound into a large roll, the additive effects of the dipole charge layers can give rise to sufficiently high potentials between separated points of the roll to cause breakdown flashes to occur in air.

The growth of the potential of the outer surface of a winding roll of film is shown in Fig. 11. A strip of 35-mm film support was wound onto a metal core, and the potential of the roll surface was measured with an electrometer and a probe. The probe was moved continuously so as to maintain a constant separation from the surface of the growing roll. Dipole charges were placed upon the film by means of the needle and roller assembly previously described. The net charge was reduced to a near zero value by directing a grounded needle toward the film in free space at a point just prior to the polarization detector, which preceded the winding roll, as shown in Fig. 11. Thus, the cumulative potential due to successive dipole layers was measured. The magnitude of the induced dipole on the film was uniform throughout the experiment.

The two curves of Fig. 11 were obtained by polarizing the film with positive and negative needle potentials, respectively. These graphs do not rise linearly from the origin as might be expected, since it was not possible to establish the desired dipole in the first few laps of film. Also, they do not rise indefinitely, but saturate at about 20 kv, a value which was not surpassed by increasing the dipole moment of the strip. This effect probably arises from leakage or discharge to air.

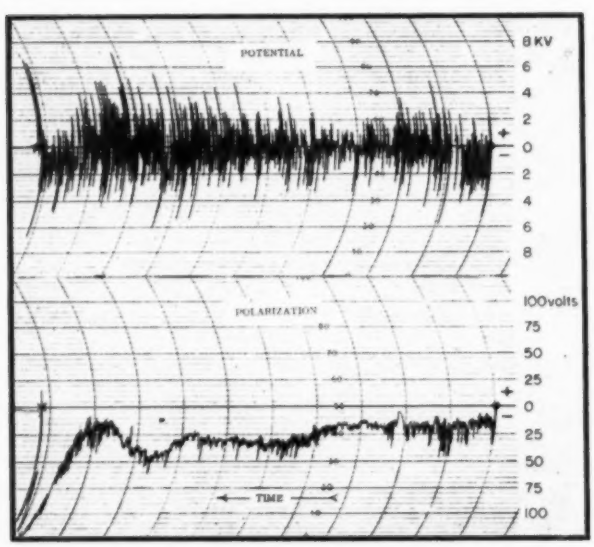


FIG. 10. Electrification records of an unwinding roll of film support.

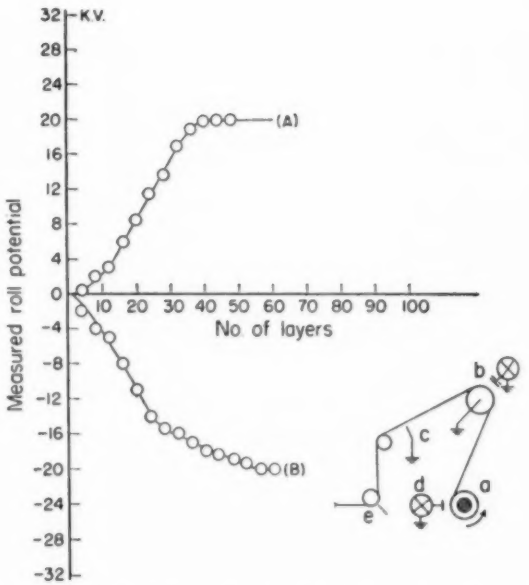


FIG. 11. Growth of film-roll potential with number of polarized film layers. (A) Needle positive; (B) Needle negative. (a) Take-up roll; (b) Polarization detector; (c) Discharging needle; (d) Electrometer; (e) Needle-roller polarizer.

The growth of the film-surface potential can be predicted from the measurement of the dipole strength of a single lap, as measured by the polarization detector. Thus, the film with the positive winding potential had a dipole potential on the grounded cylinder of 1600 volts. Consequently, the potential difference between the two faces of this filmstrip in space would be 800 volts. Therefore, one would expect the potential of the winding roll to increase at the rate of 800 volts per lap. Actually, the straight-line portion of the graph for this roll rises at the rate of 700 volts per lap. In the negative potential roll, the dipole strength of the single strip was 750 volts, and the steepest part of the curve rises at the rate of 560 volts per lap. The disparity between these figures indicates that the polarization was decaying while the measurements were being made. However, the agreement is good enough to justify the view that the potential of such a roll is the result of the addition of dipole layers.

Elastic Description of a High-Amplitude Spherical Pulse in Steel

WILLIAM A. ALLEN AND WERNER GOLDSMITH*
Michelson Laboratory, U. S. Naval Ordnance Test Station, Inyokern, China Lake, California
(Received June 25, 1954)†

Extensive calculations have been performed with an electronic calculator to evaluate a problem in elasticity that simulates the effect of a cylindrical charge of high explosive detonated in intimate contact with a steel plate. The general method of calculation has been described in detail. Although elastic theory has been extrapolated into a regime where it is known not to apply, insight of a valuable general nature has been obtained on the nature of the negative component of the pulse.

INTRODUCTION

SEVERAL papers[1–3] have appeared recently that reflect increasing general interest in the physical phenomena that occur when a charge of high explosive is detonated in intimate contact with a steel plate. The phenomena have been discussed[2] in terms of one-dimensional wave propagation in perfectly elastic plate material. Other investigators[3] have remarked that the specific geometrical arrangement of charge and specimen used in these experiments actually produces a pulse form intermediate between plane and spherically divergent. It has been shown in previous papers[4,5] that the cylindrical charge produces a spherical pulse front in the metal with much of its energy concentrated in a beam along the axis; the apparent origin of the pulse was shown to lie in the vicinity of the detonator.

It has been shown experimentally[6] that the physical effects produced near the free surface of a plate by a 30 mm diam spherical charge detonated at its center are equivalent to those produced by a standard 2-in. cylindrical charge, 1 inch in diam. The calculated pulse length and displacement amplitude produced on the free surface of a 4-in. steel specimen by a hypothetical 30 mm diam charge acting on an elastic medium is equivalent to the measured pulse length and amplitude produced by a cylindrical charge. It is concluded that further theoretical investigation of the spherically-symmetric case may throw additional light on the cylindrical charge experiment which, manifestly, is not amenable to mathematical analysis. Physical assumptions upon which the calculations are based have been discussed in the previous paper.[5]

THEORETICAL PRELIMINARIES

The partial differential equation that governs wave propagation in an elastic, homogeneous, isotropic medium can be written[7] in vector notation by means of the relation

$$\rho\frac{\partial^2\mathbf{u}}{\partial t^2}=(\lambda+2\mu)\nabla\nabla\cdot\mathbf{u}-\mu\nabla\times\nabla\times\mathbf{u}, \tag{1}$$

where λ and μ are the Lamé elastic constants[8] and $\mathbf{u}$ is particle displacement. If the divergence of Eq. (1) is taken, and the equation is divided by density ρ, one obtains the relation

$$\frac{\partial^2\nabla\cdot\mathbf{u}}{\partial t^2}=\frac{\lambda+2\mu}{\rho}\nabla^2\nabla\cdot\mathbf{u}. \tag{2}$$

Equation (2) states that the dilatation, $\nabla\cdot\mathbf{u}$, satisfies the three-dimensional wave equation. For the special case of spherical coordinates with radial displacement, Eq. (2) can be written in the form of the one-dimensional wave equation

$$\frac{\partial^2 r\nabla\cdot\mathbf{u}}{\partial t^2}=c^2\frac{\partial^2 r\nabla\cdot\mathbf{u}}{\partial r^2}, \tag{3}$$

where

$$c=\left(\frac{\lambda+2\mu}{\rho}\right)^{\frac{1}{2}} \tag{4}$$

is the velocity of dilatation waves in the medium.

It is possible to draw some general conclusions about the shape of a divergent pulse without the necessity of solving Eq. (3). A well-known result of elementary elasticity,[9] usually formulated in the classical initial value problem,[10,11] states that an initial disturbance in homogeneous isotropic media will be propagated so that at some future time t the disturbance will be bounded by a surface of two sheets. When the outer sheet b reaches any point, the medium suddenly acquires a stress and velocity, and after the inner sheet a passes the point, the medium returns to rest without stress or strain. This

* University of California, Berkeley, California.
† The original, unrevised manuscript was received on July 13, 1953.
[1] Pack, Evans, and James, Proc. Phys. Soc. (London) **60**, 1 (1948).
[2] J. S. Rinehart, J. Appl. Phys. **22**, 555 (1951).
[3] W. M. Evans and G. I. Taylor, Research (London) **5**, 502 (1952).
[4] W. A. Allen and C. L. McCrary, Rev. Sci. Instr. **24**, 165 (1952).
[5] W. A. Allen, J. Appl. Phys. **24**, 1180 (1953).
[6] W. A. Allen and W. Goldsmith, J. Appl. Phys. **25**, 813 (1954).
[7] Leigh Page, *Introduction to Theoretical Physics* (D. Van Nostrand Company, Inc., New York, 1947), second edition, p. 175.
[8] The Lamé constant μ, the modulus of elasticity in shear, is usually designated by the symbol G in engineering literature.
[9] A. E. Love, *A Treatise on the Mathematical Theory of Elasticity* (Dover Publications, New York, 1944), fourth edition, p. 301.
[10] A. G. Webster, *Partial Differential Equations of Mathematical Physics* (B. G. Teubner, Leipzig, 1933), second edition, p. 160.
[11] S. D. Poisson, Mem. de'l' Institut, t. 3 (1820).

condition can be specified for the case of divergent propagation by the symbolic relations

$$u(a)=u(b)=0, \tag{5}$$

where u represents particle displacement. Invocation of this theorem does not imply that conditions in the actual case are the same as those of the initial value problem. It seems plausible, however, that at distances great with respect to the pulse length, the conclusions from the initial value problem may be used to predict some results of the actual experiment.

It will be shown that a divergent pulse formed by an initial arbitrary radially-symmetric displacement in a homogeneous isotropic medium will be characterized by tensile components. Equation (3) signifies that the quantity $r\nabla\cdot\mathbf{u}$ moves outward from the origin with constant velocity c with no change of pulse form.[12] The dilatation $\nabla\cdot\mathbf{u}$ can be written in the form

$$\nabla\cdot\mathbf{u}=r^{-n}\partial(r^n u)/\partial r, \tag{6}$$

where $n=0, 1, 2$ for the respective cases of plane, cylindrical, and spherical propagation. Multiplication of Eq. (6) by $r^n dr$, and integration over the region of the disturbance yields the relation

$$\int_b^a r^n\nabla\cdot\mathbf{u}dr=a^n u(a)-b^n u(b). \tag{7}$$

The right-hand member of Eq. (7) will vanish for the cases of divergent propagation as a result of Eq. (5). Equation (7), in this case, reduces to the relation

$$\int_b^a r^n\nabla\cdot\mathbf{u}dr=0. \tag{8}$$

Equation (8) is interpreted to mean that dilatation $\nabla\cdot\mathbf{u}$ must change sign over the region of disturbance.[10] This result is well known both in acoustics[13] and in seismology.[14] In spherical coordinates, the dilatation can be written for radial symmetry in the form

$$\nabla\cdot\mathbf{u}=r^{-2}\partial(r^2u)/\partial r\equiv\partial u/\partial r+2u/r. \tag{9}$$

The principal stress equations can be written in the form[15]

$$\sigma_r=(\lambda+2\mu)\partial u/\partial r+2\lambda u/r, \tag{10}$$

$$\sigma_\theta=\lambda\partial u/\partial r+2(\lambda+\mu)u/r, \tag{11}$$

where σ_r and σ_θ represent, respectively, radial and tangential stress. Eliminate u_r from Eqs. (10) and (11) by means of Eq. (9) and combine to obtain the relations

$$\sigma_r=(\lambda+2\mu)\nabla\cdot\mathbf{u}-4\mu u/r, \tag{12}$$

$$\sigma_\theta=\lambda\nabla\cdot\mathbf{u}+2\mu u/r, \tag{13}$$

$$\sigma_m\equiv\tfrac{1}{3}(\sigma_r+\sigma_\theta+\sigma_\varphi)=(\lambda+\tfrac{2}{3}\mu)\nabla\cdot\mathbf{u}\equiv k\nabla\cdot\mathbf{u}. \tag{14}$$

For the cases where the second terms of the right-hand side of Eqs. (12) and (13) are negligible, the stresses assume the same sign as $\nabla\cdot\mathbf{u}$. The stresses assume the sign of the dilatation exactly in two simple cases: (1) the hydrostatic pressure of a perfect liquid; that is, when $\mu=0$ in Eqs. (12) and (13); (2) the mean stress σ_m in a solid. The constant k in Eq. (14) is defined as the modulus of volume elasticity. It can be concluded from Eqs. (3) and (14) that the quantity $r\sigma_m$ moves into the medium with constant velocity c and no change in shape.

It has been indicated that the initial value problem in the spherically-symmetric case leads to a dilatation pulse of finite length characterized by components of negative or tensile amplitude. It remains to be determined whether or not the existence of a boundary condition at the explosive-steel interface results in a pulse or a wave, and whether the assumed exponential forcing function actually simulates the interaction of the explosive and metal.

STRESS ANALYSIS

The following theoretical problem[16] will be considered: *Given a spherical cavity of radius* a *within a homogeneous ideally elastic infinite medium, find the elastic wave motion that results from application to the interior of the cavity of a pressure pulse that decays exponentially.* Elastic properties of the steel are specified by the following quantities: $E=2.139$ megabars, $\rho=7.849$ g/cc, and $\nu=0.310$ where E, ρ, and ν are Young's modulus, density, and Poisson's ratio, respectively. The pressure p imparted by the high explosive to the steel is assumed to be expressible in the form

$$p=0 \quad \text{for} \quad t<0, \tag{15}$$

$$p=p_0e^{-\alpha t} \quad \text{for} \quad t>0, \tag{16}$$

where p_0 is peak pressure, α is a time-decay constant, and time t is measured from arrival time of the detonation front at the explosive-steel interface. The properties of the high explosive are specified by the quantities: $p_0=0.283$ megabar, and $\alpha=2\,\mu\text{sec}^{-1}$. A cavity radius $a=1.5$ cm was chosen.

It can be shown[17] that spherical longitudinal pulse propagation in homogeneous isotropic media can be specified by the wave equation

$$\partial^2\varphi/\partial t^2=c^2\nabla^2\varphi, \tag{17}$$

where φ is a scalar displacement potential. A solution of Eq. (17) has been obtained[16,18] for an outgoing wave that satisfies the boundary conditions. Define the

[12] H. L. Selberg, Arkiv Fysik (Stockholm) **5**, 97 (1952).
[13] Bevan B. Baker and E. T. Copson, *The Mathematical Theory of Huygen's Principle* (Oxford University Press, London, 1950), second edition, p. 9.
[14] N. Ricker, *Proceedings of the Third World Petroleum Congress, Section I. The Hague, 1951* (E. J. Brill, Leiden, 1951), p. 514.
[15] Reference 9, p. 142.
[16] J. A. Sharpe, Geophysics, **7**, 144 (1942).
[17] Reference 9, p. 293.
[18] F. G. Blake, Jr., J. Acoust. Soc. Am. **24**, 211 (1952).

parameter

$$\tau = t - \frac{r-a}{c}. \tag{18}$$

For $\tau<0$ the potential φ is assumed to vanish and for $\tau>0$ it is specified by the relation

$$\varphi = \frac{p_0 a}{\rho r[\omega_0^2+(\alpha_0-\alpha)^2]} \times\left[-e^{-\alpha\tau}+e^{-\alpha_0\tau}\left(\cos\omega_0\tau+\frac{\alpha_0-\alpha}{\omega_0}\sin\omega_0\tau\right)\right], \tag{19}$$

where

$$\alpha_0 = \frac{c}{a}\frac{1-2\nu}{1-\nu}, \tag{20}$$

$$\omega_0 = \frac{c}{a}\frac{(1-2\nu)^{\frac{1}{2}}}{1-\nu}. \tag{21}$$

Equation (19) can be specified in a phase-angle representation[18] that leads to considerable simplification in the equations that follow. By straightforward differentiation, and collection of terms, obtain the relations

$$u = \partial\varphi/\partial r, \tag{22}$$

$$u_r = \partial u/\partial r, \tag{23}$$

$$v = \partial u/\partial t, \tag{24}$$

$$\sigma_r = (\lambda+2\mu)u_r+2\lambda u/r, \tag{25}$$

$$\sigma_\theta = \lambda u_r+2(\lambda+\mu)u/r, \tag{26}$$

$$\sigma_m = (\lambda+\tfrac{2}{3}\mu)(u_r+2u/r), \tag{27}$$

$$\sigma_s = \mu(u_r-u/r), \tag{28}$$

where σ_s is defined as maximum shear stress.

The calculations associated with the preceding equations were performed by means of an IBM card programmed electronic calculator which has been described[19] previously in connection with the ray tracing problem of geometrical optics. The floating decimal formulation described in this previous paper was found to be useful in the present problem. All calculations were carried to nominal eight-figure accuracy. The optical problem can be formulated in a way that avoids the use of elementary functions usually obtained from tables. In the present problem, however, the machine calculates, without recourse to tabulated data, values of $\sin x$, $\cos x$, $\sin^{-1}x$, and e^x whenever these quantities are needed. The stress Eqs. (25) to (28) were solved for a network of values $\tau=0, 1, 2, \cdots, m$, and $r=a+cl$, where $l=0, 1, 2, \cdots, n$. The displacement potential φ was used by the calculator in its phase angle representation which requires calculation of an inverse trigonometric function. It might be informative to list the equations used by the calculator in the same form and in the same sequence in which they were solved:

$$\left.\begin{matrix}E\\ \rho\\ \nu\end{matrix}\right\}\text{material constants} \qquad \left.\begin{matrix}p_0\\ \alpha\\ a\end{matrix}\right\}\text{explosive constants}$$

$$\mu = \tfrac{1}{2}E/(1+\nu), \tag{29}$$

$$h^2 = 1-2\nu, \tag{30}$$

$$\lambda = 2\mu\nu/h^2, \tag{31}$$

$$c^2 = (\lambda+2\mu)/\rho, \tag{32}$$

$$\omega_0 = ch/[a(1-\nu)], \tag{33}$$

$$\alpha_0 = h\omega_0, \tag{34}$$

$$A^2 = \omega_0^2+(\alpha_0-\alpha)^2, \tag{35}$$

$$\beta = \sin^{-1}(\alpha_0-\alpha/A), \tag{36}$$

$$B = p_0 a/\rho A^2\omega_0, \tag{37}$$

$$C = \omega_0 e^{-\alpha\tau}, \tag{38}$$

$$D = Ae^{-\alpha_0\tau}\cos(\omega_0\tau-\beta), \tag{39}$$

$$H = A\omega_0 e^{-\alpha_0\tau}\sin(\omega_0\tau-\beta), \tag{40}$$

$$r\varphi = B(D-C), \tag{41}$$

$$F = -\alpha C+\alpha_0 D+H, \tag{42}$$

$$G = \alpha^2 C+(\omega_0^2-\alpha_0^2)D-2\alpha_0 H, \tag{43}$$

$$u_1 = Bc^{-1}F, \tag{44}$$

$$u_2 = -r\varphi, \tag{45}$$

$$v_1 = Bc^{-1}G, \tag{46}$$

$$v_2 = BF, \tag{47}$$

$$w_1 = -c^{-1}v_1, \tag{48}$$

$$w_2 = -2u_1, \tag{49}$$

$$w_3 = -2u_2, \tag{50}$$

$$u = u_1 r^{-1}+u_2 r^{-2}, \tag{51}$$

$$v = v_1 r^{-1}+v_2 r^{-2}, \tag{52}$$

$$w \equiv u_r = w_1 r^{-1}+w_2 r^{-2}+w_3 r^{-3}. \tag{53}$$

The stresses are calculated at this point by means of Eqs. (25) to (28). All calculations are facilitated by means of a proper choice of units. The cm, g, and μsec are the natural units to use in explosion research. Density ρ retains its normal cgs value in this system, and quantities with units of pressure, such as E, p_0, λ, and μ, are expressed in megabars. The megabar is used commonly to express the magnitude of pressure produced by explosions.[5,20,21] Macroscopic velocities produced in ordnance research range in value up to 9 cm/μsec ob-

[19] W. A. Allen and R. N. Stark, J. Opt. Soc. Am. **41**, 636 (1951).

[20] Feynman, Metropolis, and Teller, Phys. Rev. **75**, 1561 (1949).

[21] Walsh, Shreffler, and Willig, J. Appl. Phys. **24**, 349 (1953).

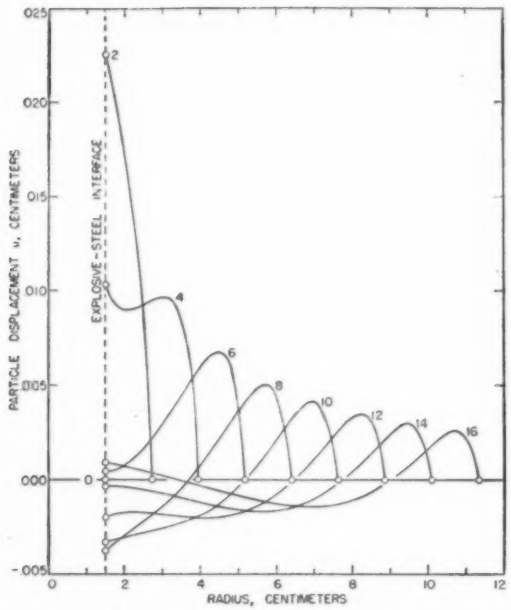


FIG. 1. Particle displacement as a function of radius. The numbers on the curves denote time in μsec.

tained for shaped charge jets.[22] Sound velocity in these units amounts to 0.6 cm/μsec for steel.

RESULTS

Results obtained from the calculation procedure are illustrated in Fig. 1 to Fig. 5 where particle displacement u, particle velocity v, radial stress σ_r, and tangential stress σ_θ are plotted for times $t=0, 2, 4, \cdots, 16$ μsec. The origin of time is taken as arrival time of the detonation front at the explosive-steel interface. Particle displacement curves are plotted in Fig. 1. The maximum theoretical elastic particle displacement is less than 0.024 cm. At about 4 μsec, the displacement curve begins to assume the character of a pulse. At about 7 μsec, the cavity begins to decrease in diam below its original value of 1.5 cm. For times 8, 10, and 12 μsec, most of the disturbed radius exhibits a negative displacement, with a whiplike motion of the interface end of the curve. At about 15 μsec, the cavity radius again begins to increase in excess of 1.5 cm. The cavity will oscillate with attenuating amplitude and increasing period for all time.

Particle velocity curves are plotted in Fig. 2. The maximum particle velocity predicted by elastic theory, 0.059 cm/μsec, compares favorably with the value 0.069 cm/μsec predicted by English investigators[1] based on an entirely different, but more exact, theory. The initial effect rapidly becomes effectively a sharp positive pulse of short length followed by a pronounced negative component.

Radial stress curves are plotted in Fig. 3. As a result of the assumed exponential behavior of the pressure pulse in the cavity, the pressure rapidly approaches a negligible value at the interface. At about 3 μsec, the pulse has developed a pronounced negative component.

Tangential stress curves are plotted in Fig. 4. Note the violent stress reversal at the interface and the relative prominence of the tensile component compared with the compressive component of the wave.

It can be inferred from Fig. 3 and Fig. 4 that the formula $\sigma=\rho cv$ does not apply in general. Stress σ_r essentially vanishes over the interface at $t=4$, but particle velocity v, Fig. 2, remains large. At the wave front $\tau=0$, however, the following relations can be verified from Eqs. (29) to (53):

$$u=0, \tag{54}$$

$$v=p_0a/\rho cr, \tag{55}$$

$$\sigma_r=p_0a/r=\rho cv, \tag{56}$$

$$\sigma_\theta=p_0a\lambda/(\lambda+2\mu)r=\rho cv\lambda/(\lambda+2\mu). \tag{57}$$

Equations (55), (56), and (57) imply that peak particle velocity v and peak stresses σ_r and σ_θ attenuate at a rate inversely proportional to radius r. However, attenuation at points behind the wave front is described by additional components varying inversely with higher powers of the radius, Eqs. (52) and (53).

The initial value problem states that a pulse that consists of positive and negative components can traverse the medium leaving undisturbed material behind it. On the other hand, the radial stress curves of Fig. 3 indicate that the assumed conditions of the problem

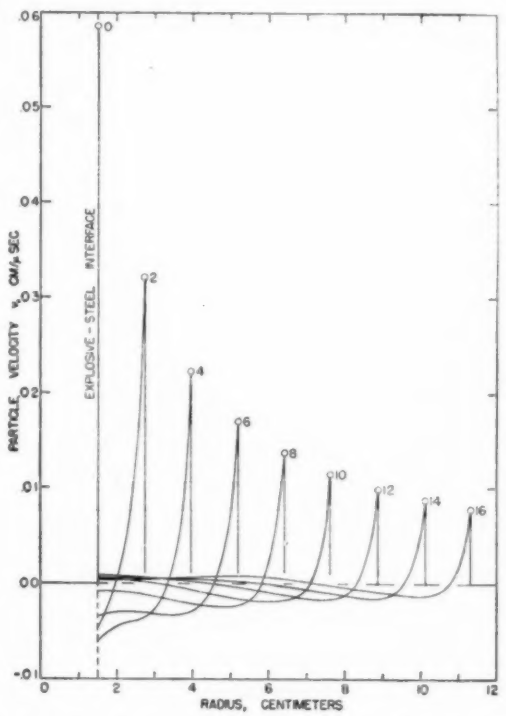


FIG. 2. Particle velocity as a function of radius. The numbers on the curves denote time in μsec.

[22] Koski, Lucy, Shreffler, and Willig, J. Appl. Phys. **23**, 1300 (1952).

requires the existence of a damped sinusoidal wave that extends from the interface to the pulse front. Thus, the assumptions upon which the initial value problem is based do not obtain, and the arguments in the section on theoretical preliminaries are not necessarily pertinent. The sinusoidal wave attenuates rapidly, however, and the resultant effect is essentially that of a pulse of constant length. Experimental data[5,14] seems to confirm that a pulse of definite length is actually produced. Figure 3 and Fig. 4 suggest the enormous tensile forces necessary for stability. Curves representing the shearing stresses, not plotted, show similar stress reversal.

The plotted curves are not exhaustive. Curves for other initial conditions with different values of a and α have been compiled. It would also have been possible to

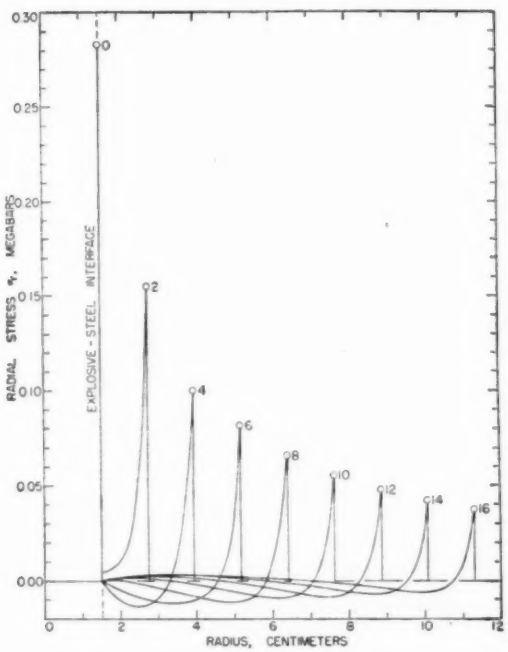


FIG. 3. Radial stress as a function of radius. The numbers on the curves denote time in μsec.

amplify the graphs by plots of maximum shear stress. It was decided, however, that equivalent information could be provided by means of one additional graph. Figure 5 represents a plot of the quantity $r\sigma_m$ which is known not to change shape as it moves into the medium. The plot was made for $t=15$ μsec. The quantity $r\sigma_m$ can be obtained for all times and for all distances by means of a simple translation of the distance axis, and Fig. 5 thus represents a single comprehensive diagram representing the stresses associated with this phenomenon.

Figure 3 and Fig. 4 indicate that the negative component associated with σ_θ is considerably greater than that associated with σ_r in the vicinity of the cavity. This condition will give rise to large shear stresses in this region.

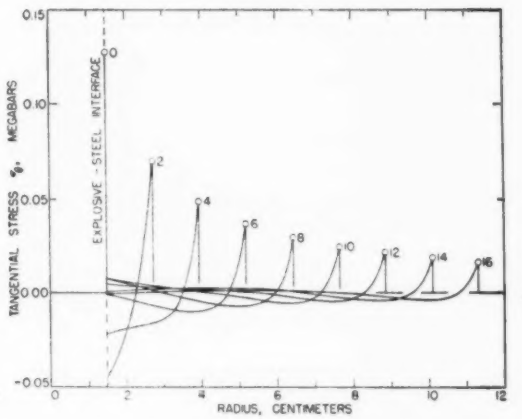


FIG. 4. Tangential stress as a function of radius. The numbers on the curves denote time in μsec.

CONCLUSIONS

The calculated stresses in this problem greatly exceed the static tensile strengths of all low and medium carbon steels. The ultimate tensile strength of these materials, however, may be raised considerably under the extremely rapid loading conditions existing in the experiment. A more plausible explanation of the inordinate stress magnitudes lies in the extrapolation of elementary elastic theory into a physical regime where the theory is known not to apply. The exponential form of the forcing function, Eq. (16) was assumed for mathematical convenience; the function probably does not accurately represent the actual situation. Experimental[2] and theoretical[23] evidence indicate that the initial pulse in the high explosive has a more complicated form. No attempt is made to estimate the effect on the pulse form of interaction at the explosive-steel interface.

Although the actual problem does not fulfill the conditions of the initial value problem owing to the existence of the boundary condition, it seems plausible that some of the predictions of the initial value problem will still hold, especially after the pulse travels a distance large

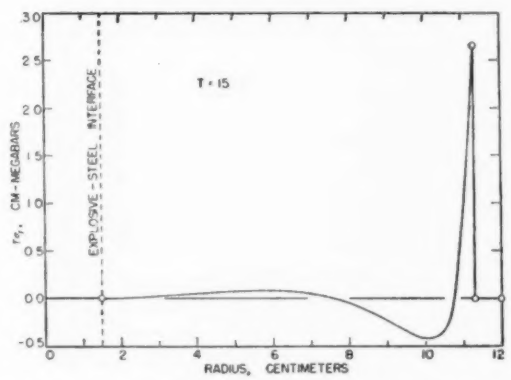


FIG. 5. Product of mean stress and radius measured from the origin.

[23] R. Hill and D. C. Pack, Proc. Roy. Soc. (London) **191**, 524 (1947).

with respect to the diameter a. It may be inferred from stability arguments that the elastic pulse must have both positive and negative components. The initial positive pulse must rapidly develop a negative phase in order to exist in divergent propagation. The pulse form measured by a previous investigator[2] probably has a negative phase which could not be measured by the Hopkinson bar method employed. Experiment[5] and theory[14,24] both lead to a quasi-pulse that leaves the medium relatively undisturbed after its passage. Negative pressure components are pronounced, expecially for the case of tangential stress.

[24] S. Homma, Geophys. Mag. **23**, 145 (1952).

ACKNOWLEDGMENT

It is a pleasure to acknowledge the technical assistance of Mr. J. W. Rogers who made the illustrations, and Mr. Maurice Aegerter of the Computing Branch who coded the problem for the IBM card programmed electronic calculator.

JOURNAL OF APPLIED PHYSICS VOLUME 26, NUMBER 1 JANUARY, 1955

Relaxation Oscillations and Noise from Low-Current Arc Discharges

M. I. Skolnik* and H. R. Puckett, Jr.†
Radiation Laboratory, The Johns Hopkins University, Baltimore, Maryland
(Received April 26, 1954)

Under certain conditions the usual gas-discharge circuit was found to give rise to relaxation oscillations. The relaxation oscillations produce a series of very narrow pulses with random pulse repetition interval, which appear as noise. The amount of noise was found to depend upon the gas, the cathode electrode material, the power supply voltage, and the external circuit configuration. It seems that this mechanism may be responsible for most of the high level noise usually reported from gas discharges. The pulsed nature of the apparently continuous discharge appears to be a fundamental property of the low-current, cold-cathode arc which has not been considered previously.

INTRODUCTION

MANY examples have been reported of gas discharges which produce electrical noise in the rf region of the electromagnetic spectrum. Noise has been found from thyratrons,[1] corona discharges,[2] glow discharges,[3,4] and fluorescent lamps.[5] Relaxation oscillations have also been observed with gas discharges. In this paper, the relaxation oscillations and the noise obtained with low-current electric arcs will be described. The noise has been detected over a broad range of frequencies, extending from low frequencies up to the microwave region. It is shown that the noise is a result of the relaxation oscillations and is dependent upon the circuit parameters as well as the nature of the gas discharge.

NOISE MEASUREMENTS

A convenient parameter with which to describe the noise power is the *noise factor*. This may be defined as the ratio of the noise power of the device to the available noise power from a resistor at room temperature, when measured over the same band width. The available noise power over a band width B from a resistor at an absolute temperature T is $P=kTB$ where k is the Boltzmann constant.[6] At room temperature (293°K) the available noise power from a resistor is 4×10^{-21} watt/cycle. For convenience, the noise factor is usually expressed in decibels.

The noise factor of low-current electric arcs was measured over a frequency range from 0.1 mc to 4500 mc using conventional superheterodyne receivers. A substitution technique was employed. The noise power was indicated by a dc meter in the second detector circuit of the receiver. A cw signal of known, but variable, amplitude was substituted and adjusted until the second detector meter again gave the same reading. The cw power divided by the receiver band width is equal to the noise power per unit band width. The ratio of the noise power per unit cycle of band width to 4×10^{-21} is the noise factor. The substitution technique assumes that the second detector meter responds equally to noise signals and cw signals of the same power.

All measurements were carried out with the arc in a demountable metal chamber as shown in outline in Fig. 1. Although the chamber was not necessary for the atmospheric pressure arcs in air, it was used with all gases tested so that the stray circuit capacities were the same whether air or some other gas were used.

* Now with M. I. T. Lincoln Laboratory, Lexington, Mass.
† Presently serving on active duty with the U. S. Army, Corps of Engineers.

[1] J. D. Cobine and C. J. Gallagher, J. Franklin Inst. **243**, 41–54 (1947).
[2] R. L. Tanner, Stanford Research Institute, Tech. Report No. 37, (April, 1953).
[3] N. R. Labrum and E. K. Bigg, Proc. Phys. Soc. (London) **B65**, 356–368 (1952).
[4] S. Ruthberg, NDRC Report No. OEMsr–411–239 (September 13, 1945).
[5] W. W. Mumford, Bell System Tech. J. **28**, 698–618 (1949).
[6] H. Nyquist, Phys. Rev. **32**, 110–113 (1928).

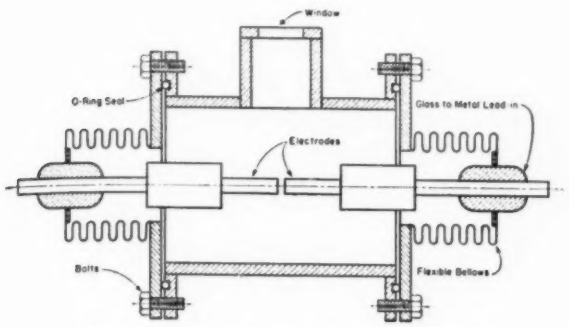


FIG. 1. Demountable gas-discharge chamber.

This procedure was necessary since the results obtained were affected by the amount of stray capacity across the gas discharge.

The basic circuit diagram is shown in Fig. 2. The capacity C need not necessarily be a physical circuit element, but may be the stray capacity of the circuit. The rf noise was coupled from a series resistor in the cathode circuit of the gas discharge into a section of coaxial cable. The output was taken from the cathode lead since it was found that slightly more noise was obtained than when the output was taken from the anode lead. All circuit leads were as short as possible and no attempt was made to compensate the rf coupling circuit over the frequency range.

In Fig. 3, is plotted the noise factor in decibels as a function of frequency for $\frac{1}{4}$ inch diameter aluminum electrodes in air at atmospheric pressure. The current as read by a dc meter was 50 ma and the electrode separation was 0.01 inch. The points plotted in the figure represent the experimental measurements while the dotted curve is a computed spectrum and will be discussed in a later section. As a comparison the 884 thyratron noise generator[1] has a maximum noise factor of approximately 114 db at a frequency of 700 kc and the fluorescent tube noise generator has a noise factor of 15.5 db at 3000 mc.[7] Labrum and Bigg[3] report noise factors of 85 db at approximately 200 mc with low-pressure glow discharges in air.

RELAXATION OSCILLATIONS

When an oscilloscope was placed across the resistor r in Fig. 2, a series of pulses as shown in Fig. 4 was obtained.

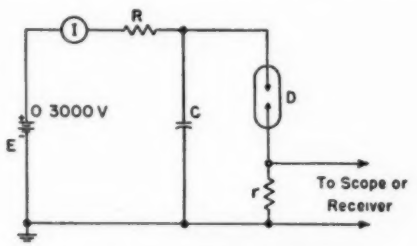


FIG. 2. Gas-discharge circuit.

[7] H. Johnson and K. R. DeRemer, Proc. Inst. Radio Engrs. 39, 908-914 (1951).

The exposure time of this figure was 1/25 second, and many sets of pulses were superimposed. The initial pulse appears narrower than the succeeding pulses because the scope triggered so that the initial pulse of each sweep occurred at the same spot but the succeeding pulses coincided only if the pulse repetition interval remained constant. The width of the succeeding pulses is, therefore, a measure of the randomness of the pulse repetition interval. An individual pulse is shown on an expanded time scale in Fig. 5. The pulse was obtained by applying the output across the resistor r of Fig. 2 directly to the plates of a Tektronix 517 oscilloscope (7 millimicroseconds rise time). The duration of the pulse is rather short, approximately 25 millimicroseconds, and represents the discharge‡ of the capacitor C.

It was found that the pulses were a consequence of the relaxation oscillator circuit formed by the series limiting resistor and the capacity across the gas discharge. It was not realized at the outset of these experiments that the circuitry associated with the gas-discharge was of the type shown in Fig. 2 since no physical capacitor C was deliberately placed across the gas discharge. Instead, the capacity C was the stray circuit capacity. The stray capacity of the circuit configuration was found to be about 25 micromicrofarads which gave rise to a relaxation frequency of about 4.5 mc when R was 15 000 ohms. Since it was more convenient for measurement purposes to operate the discharge with a lower repetition frequency, a 150 micromicrofarad capacitor was placed across the tube from anode to ground, making the total shunt capacity approximately 175 micromicrofarads. This reduced the repetition frequency to 650 kc. The stray capacity was associated primarily with the circuit elements and the connecting wiring. The capacity formed by the electrodes of the gas discharge was negligible in comparison.

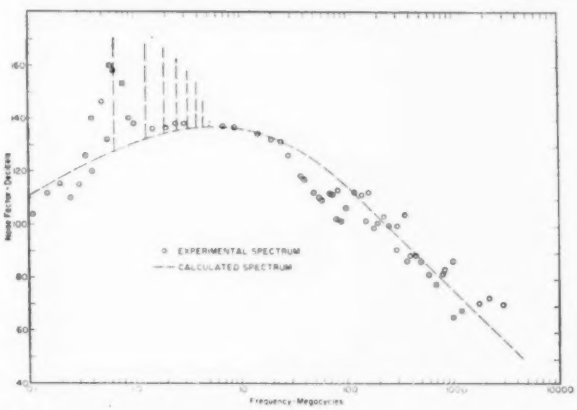


FIG. 3. Measured and computed spectrum of aluminum arc discharge operated in air at atmospheric pressure, with $\frac{1}{4}$ inch diameter rod electrodes, 0.01 inch electrode separation, and a current of 50 ma.

‡ Since the term "discharge" can apply to either the electric arc or to the process whereby the capacitor loses its charge, in the remainder of this paper "gas discharge" will be used to indicate the former and "discharge" will indicate the latter.

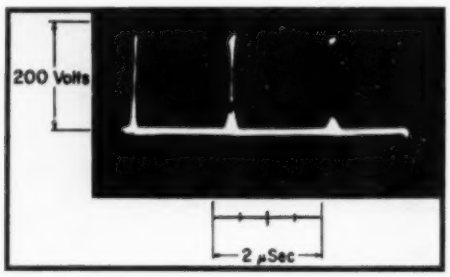


FIG. 4. Series of pulses formed by gas-discharge circuit for same experimental conditions as Fig. 3.

The period of relaxation oscillations is given by Eq. (1),

$$T = RC \log_e \frac{E - E_0}{E - E_i}, \tag{1}$$

where R, C, and E are defined in Fig. 2 and E_0 is the voltage at which the gas discharge is extinguished and E_i is the voltage at which gas discharge is initiated.[8] The validity of Eq. (1) was tested by varying R, C, and E and observing the corresponding changes in T. The equation was found to fit the measured value of T to within the limit of accuracy with which one can measure distances on the face of a cathode ray tube with a display as shown in Fig. 4. If T, R, and C are known for two different values of E, then E_0 and E_i may be found. Because of the exponential relationship, however, relatively large errors in the determination of these voltages can occur with only small errors in the measurement of T. Hence, E_0 and E_i cannot be found from Eq. (1) with any degree of accuracy, unless T/RC is accurately known. It appears, however, that as best as can be determined, E_0 is of the order of 10 to 20 volts and E_i is about 450 to 500 volts.

COMPUTED SPECTRUM

It will now be shown that the pulses formed by the relaxation oscillations described above can account for the noise observed. Because the repetition rate is not uniform, the spectrum consists of a continuous portion as well as a discrete portion. The continuous spectrum is the "noise."

Lawson and Uhlenbeck[9] give an expression,

$$\langle G(f)\rangle_{Av} = \frac{2}{\theta_0}|B(f)|^2[1-|\phi(f)|^2] + \frac{2}{\theta_0^2}|B(f)|^2|\phi(f)|^2\sum_{n=0}^{\infty}\delta\left(f-\frac{n}{\theta_0}\right) \tag{2}$$

for the power spectrum $\langle G(f)\rangle_{Av}$ of a series of pulses that have identical shape and height but a repetition period varying around an average value according to some probability distribution. This expression can be used to compute a theoretical spectrum from the information given in Figs. 4 and 5 for the pulse repetition characteristics and the pulse shape. The average repetition period of the series of pulses is θ_0 and the deviation of the kth spacing from θ_0 is given by e_k. The average value of e_k is assumed zero. If $P(e)$ is the probability distribution for e, then $\phi(f)$ is given by Eq. (3):

$$\phi(f) = \int_{-\infty}^{\infty} \exp(2\pi i f e) P(e) de. \tag{3}$$

Assuming a Gaussian distribution for $P(e)$ with standard deviation of σ, gives $\phi(f)$ equal to $\exp(-2\pi^2\sigma^2 f^2)$. The frequency spectrum of a single pulse is $B(f)$ and is defined in Eq. (4),

$$B(f) = \int_{-\infty}^{\infty} F(t) \exp(-2\pi i f t) dt, \tag{4}$$

where $F(t)$ describes the shape of the pulse.

The pulse shown in Fig. 5 can be approximated by Eq. (5),

$$F(t) = 9\times10^{10} t \exp(-1.8\times10^{8} t), \tag{5}$$

where t is in seconds and $F(t)$ is in volts. For the particular set of circuit parameters chosen, the form of Eq. (5) shows that the pulse was almost critically damped. The shape of the pulses was quite uniform for a given set of experimental conditions. The average pulse period θ_0 and the deviation σ was determined from *CRO* displays similar to those of Fig. 4. For the particular case of the aluminum arc in air, θ_0 was 1.54×10^{-4} sec, while σ was estimated to be about 5 percent of θ_0. The value of σ cannot be determined to a high degree of accuracy from these CRO displays, but this is not too important since a high degree of accuracy is not necessary in order to remain within the limits of measurement errors.

The presence of the δ function in the second term of Eq. (2) gives rise to a discrete spectrum in addition to the continuous spectrum given by the first term. There will be a component of the discrete spectrum whenever

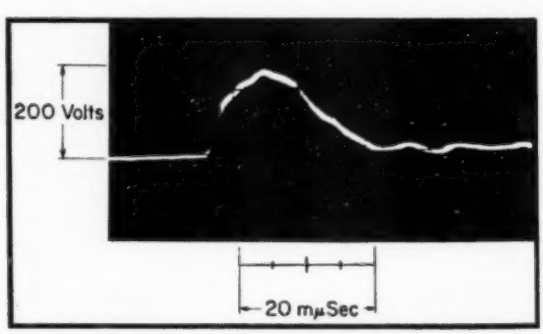


FIG. 5. Individual discharge pulse. Same experimental conditions as Fig. 3.

[8] S. Seely, *Electron-Tube Circuits* (McGraw-Hill Book Company, Inc., New York, 1950), p. 444.

[9] J. L. Lawson and G. E. Uhlenbeck, *Threshold Signals*, MIT Radiation Laboratory Series (McGraw-Hill Book Company, Inc., New York, 1950), p. 44.

$f=n/\theta_0$, where n is an integer. The amplitude of the discrete spectrum is determined by the nature of the function $\phi(f)$ as well as by the shape of a single pulse.

The power spectrum as given by Eq. (2), must be divided by the resistance r, (Fig. 2), across which it is measured in order for the continuous part of the computed spectrum to be given in watts per cycle and for the discrete part to be given in watts. To convert the continuous part of the spectrum to noise factor, the number of watts per cycle is divided by 4×10^{-21}. Although it is not appropriate to speak of a noise factor for the discrete portion of the spectrum, the amplitudes of the discrete frequency components were divided by kTB and plotted as noise factor. This procedure allowed a comparison to be made between the measured and the computed spectrum. The band width B was the band width of the measuring apparatus (3 kc over the discrete spectrum frequency range).

The computed noise power spectrum for the aluminum arc in air is shown in Fig. 3 by the dotted curve. The fit of the computed spectrum to the experimental points seems satisfactory when it is considered that (1) the method for estimating the deviation of e is not too accurate; (2) the analytical expression for the shape of an individual pulse does not include the fine structure shown in Fig. 5; (3) the use of different measuring equipments to cover the entire frequency range caused changes in the shape of the pulse because of the different input circuits associated with each type equipment; (4) the short circuiting of the arc whenever material was transferred from one electrode to another caused instabilities and made the measurement of average noise level difficult at times; and (5) the model of a Gaussian distribution may not properly describe the exact nature of the randomness of the pulse repetition interval.

Although no measured values of noise factor are shown in Fig. 3 which correspond to the discrete spectrum, other than the fundamental, their presence was recognized. In the frequency range in which the discrete spectrum was prominent, higher values of noise factor could be measured than shown in Fig. 3, but these measurements were very unstable and were not recorded. The instability was probably due to the combined effects of a narrow measuring band width (3 kc) and some variation in the average pulse repetition rate.

The agreement between the measured noise spectrum and the computed spectrum indicates that the noise detected from the low-current arc in air can be attributed to relaxation oscillations whose repetition interval is random. The height of the pulses depends on the initiation voltage and the extinguishing voltage provided the circuit through which the capacitor discharges is not changed.

Experimentally it was found that the noise power and the pulse repetition frequency increased almost linearly with an increase in gas-discharge current. This was caused by the increased power supply voltage required to increase the current. An increase in power supply voltage causes a decrease in the average pulse repetition interval [$dT/dE<0$, Eq. (1)], and hence more average noise power. The average current in each pulse can be approximated by the product of the average dc current (50 ma) times the duty ratio of the pulses ($1.54\times10^{-6}/2.0\times10^{-3}$) which yields an average pulse current of 3.85 amperes.

TABLE I.

Electrode	SF_6	O_2	Air	N_2	A	He
Aluminum	*A*–1	*A*–1	*A*–1	*A*–3	*A*–3	*B*–2$^+$
Zinc	*A*–2	*A*–2	*A*–3	*C*–3	*C*–3	*D*
Lead	*A*–2$^+$	*C*–3	*C*–1$^+$	*B*–3	*C*–3$^+$	*D**
Copper	*A*–2	*A*–1$^+$	*C*–3	*D*	*D**	*D**
Tin	*A*–3	*B*–2	*B*–3	*C*–3$^+$	*D**	*D**
Thoriated *W*	*A*–1	*D**	*C*–3	*B*–2$^+$	*B*–2$^+$	*D**
Tungsten	*A*–1	*D*	*D*	*B*–3$^+$	*D**	*D**
Molybdenum	*A*–1$^+$	*D**	*D*	*B*–3$^+$	*A*–1	*D**
Silver	*A*–3	*C*–3	*C*–1	*C*–3	*C*–3	*D**
Carbon	*A*–2	*D**	*C*–1$^+$	*C*–3$^+$	*C*–2	*D**
Nickel	*A*–1	*D*	*D*	*D**	*C*–3	*D**
Tantalum	*A*–2$^+$	*B*–2	*D*	*D*	*C*–3	*D**
Mild Steel	*A*–2	*D**	*D*	*D*	*C*–3	*D**
Titanium	*B*–3	*D*	*D*	*D**	*D*	*D**

Code: *A*—very strong, noise factor >114 db (measured at 50 mc). *B*—strong, noise factor: 114–84 db. *C*—weak, noise factor: 84–50 db. *D*—negligible or no noise. 1—steady noise level. 2—fluctuating noise level. 3—very unsteady noise level. +—combinations with very random pulse repetition interval. *—low amplitude oscillations observed.

OTHER ELECTRODES AND GASES

In Table I is listed the qualitative results that were obtained with various electrode materials and gases. In general, the gases are listed from left to right in order of decreasing noise level, and the electrode materials are listed in the same manner from top to bottom. The letter listed for each combination is a measure of the noise factor at a frequency of 50 mc while the number associated with each letter indicates the degree of short-term stability. The discharge was usually allowed to stabilize for at least five minutes before measurements were recorded since the noise did not remain at its initial level with certain materials and gases, but could change during the first few minutes of operation. When the noise factor changed without being caused by a change in any of the external parameters, it was usually due to a spontaneous change in the pulse repetition interval rather than a change in the pulse amplitude. With materials such as tin and lead, time was not available for the discharge to operate a full five minutes before measurements were made because these materials melted too rapidly.

All tests were carried out at atmospheric pressure in the demountable chamber shown in Fig. 1. No special precautions were taken to insure absolutely pure experimental conditions. All materials were used as received from the manufacturer except that the electrode materials were cleaned with sandpaper. Whenever practical, the electrode materials were used in the form

of $\frac{1}{4}$-inch diameter rods. Exceptions to this were tin, silver, and zinc which were used as $\frac{1}{16}$-inch diameter rods. Nickel and tantalum were available only in sheet form and were used in $\frac{1}{4}$-inch wide strips, 20 mils thick, and bent at one end so as to form a $\frac{1}{4}$-inch square face.

Whenever possible, all measurements were made with a power supply voltage of 1000 volts. The average current was from 40 to 50 ma. If 1000 volts was not high enough to maintain the gas discharge, the voltage was raised until satisfactory operation was obtained. When the voltage had to be greater than 1000 volts, it was usually adjusted to produce a current of 50 ma. The electrode separation was approximately 0.01 inch. However, it was found that changes in electrode separation from 0.005 inch to 0.02 inch had little effect on the amount of noise observed.

The pulses that were formed appeared similar to the pulse shown in Fig. 5 for the air-Al arc. The height of the pulse depended upon the type of gas and electrode but the pulse duration was essentially independent of these parameters. The degree of randomness of the pulse repetition interval was found to be very large in some cases. The entries in Table I marked with a (+) had a particularly large dispersion of the pulse repetition interval.

When the gas-discharge pulses were produced (relaxation oscillations), the appearance of the gas discharge was characteristic of cold-cathode arcs. The arc column was narrow and the small cathode spot appeared to wander rapidly over the face of the electrode. This was true even for tungsten electrodes whenever they produced pulses. When the arc did not operate so as to produce pulses, the cathode spot was stationary and much larger than with the cold-cathode arcs, although the anode spot was small. When no noise was observed, the discharge had the general appearance of the quiet, stable, thermionic arc. The pulse mechanism appears to be a characteristic of those gas discharges usually known as low-current, cold-cathode arcs.

In order to determine whether the noise mechanism was associated to a greater degree with either the anode or the cathode, several combinations of electrode materials were tried with one another. It was found that the noise produced by the gas discharge was characteristic primarily of the cathode material. The anode material had little effect so long as it was not deposited on the cathode electrode. This information combined with the observation that more noise was obtained when the output was taken from the cathode electrode indicates that the nature of the cathode is very important.

In some, but not all, cases of arcs which did not produce pulses, low level oscillations of a fraction of a volt in amplitude and from one to ten megacycles in frequency were observed. These cases are indicated with an asterisk (*) in Table I. The oscillations appear to be different from the relaxation oscillations not only because of the large differences in amplitudes, but also because the amplitude and frequency were dependent upon the discharge current and the electrode spacing. The frequency was not affected by the circuit capacity. The nature of these oscillations was not definitely established, but they may possibly be due to plasma oscillations of the positive ions.

In chemically active gases, deposits of compounds were usually formed on the electrodes. It seemed that those electrode-gas combinations which reacted chemically with one another produced more noise than those which did not. It is not known, however, whether this is a causal relationship or not.

It should be cautioned that the results presented in Table I were obtained with electrodes and gases which received no special treatment.

NOISE FROM OTHER DISCHARGES

Gas discharges are generally thought to be noisy devices. There appear to be three basic mechanisms which may be responsible for the production of noise in gas-discharges: (1) thermal noise, (2) shot noise, and (3) noise resulting from relaxation oscillations. Both thermal noise and shot noise are of relatively low magnitude and are usually no greater than 20–25 db noise factor. It seems likely that any larger amount of noise which may be found from gas discharges may be caused by relaxation oscillations. The presence of relaxation oscillations may not be suspected in the usual type of gas discharge, since physically placing a capacitor circuit element across the gas-discharge tube is not required. The stray capacity associated with the circuit may be as high as a hundred or more micromicrofarads and this stray capacity in conjunction with the series limiting resistor can form the relaxation oscillator circuit. The statistical factors involved in the breakdown can probably introduce the randomness necessary to produce a continuous spectrum.

The general shape of the noise spectra produced by an 884 thyratron[1] and by cold-cathode discharges,[4] seems to be of the type that could be accounted for by a relaxation oscillator mechanism. The noise reported by Labrum and Bigg[3] from glow discharges may also be of this type. Although the noise output from a fluorescent tube has been shown to be primarily due to a thermal mechanism,[10] low-frequency fluctuations from these tubes have been reported.[7] The fluctuations from a fluorescent tube have also been observed on a cathode-ray oscilloscope by the authors. Another instance in which noise can probably be attributed to the pulses resulting from relaxation oscillations is the corona discharge. The fact that noise is produced by corona discharges is well documented[2] and the pulsed nature of corona discharges has also been well established.[11] The electrical noise produced by lightning discharges is an example of a case similar to the pulsed mechanism dis-

[10] P. Parzen and L. Goldstein, Phys. Rev. **82**, 724–726 (1951).
[11] L. B. Loeb, J. Appl. Phys. **19**, 882–897 (1948).

cussed.[12] Over a large geographical area, there will be many random strokes which will add to give a continuous spectrum resulting in radio atmospheric noise.

DISCUSSION

It has been shown that the short-duration pulses which are due to the discharge of a capacitor through a gas discharge can account for some types of noise found with gas discharges. The fundamental mechanism is that of a relaxation oscillator. The amount of noise produced depends upon the gas and the electrode material, as well as the external circuit configuration. The frequency of the recurring pulses depends upon the magnitude of the series limiting resistor, the capacity across the gas-discharge device, the power supply voltage, the breakdown voltage of the gap, and the extinguishing voltage of the gas discharge. The peak current through the gas discharge can be several amperes and the form of the gas discharge is a recurring electric arc of short duration. The width of the pulse is determined by either the rc circuit through which the capacitor discharges or by the buildup time of the gas discharge, whichever is the longer. The results seem to indicate that the pulse width was limited chiefly by the discharge circuit for those cases for which substantial noise was found, since the pulse widths were approximately the same in all those cases. The pulse repetition interval is determined by the parameters of Eq. (1), provided the gas-discharge gap can deionize within this time.

Table I indicates that the electrode material as well as the gas determines the amount of noise produced. The presence of electronegative gases, however, seems to be important since such gases require larger breakdown voltages than gases which do not form negative ions.[13] Electronegative gases are also important for the rapid quenching (deionization) of gas discharges.[14] The former property results in a larger pulse amplitude while the latter yields a shorter pulse. Both conditions yield increased noise. The nature of the cathode material and its surface condition was also found to be a very important part of the noise mechanism. Further experiments under more controlled conditions appear desirable in order to determine more precisely the parameters involved.

ACKNOWLEDGMENT

The authors wish to acknowledge the assistance of Mr. Nathaniel Perry during the experimental phase of this work.

[12] B. F. J. Schonland, *The Flight of Thunderbolts* (Oxford University Press, London, 1950), p. 136.

[13] G. Camilli and R. E. Plump, Trans. Am. Inst. Elec. Engrs. Communications and Electronics, No. 6, 93–102 (May, 1953).

[14] F. Llewellyn Jones, Proc. Inst. Elec. Engrs. (London) I. **96**, 305–312 (1949).

An Approximate Theory of Armor Penetration*

WILLIAM T. THOMSON †
University of California, Los Angeles, California
(Received March 29, 1954)

The problem of armor penetration of thin plates is considered from a quasi-dynamical approach. Equations are derived for the energy dissipation due to plastic deformation and for heating of the projectile target interface. Both the conical and the ogival head are considered in the application of the general equations.

THE problem of armor penetration is a complex one, and by its military nature much of the literature on the subject is classified. The analysis of the problem requires the making of many idealized approximations which in turn define the problem to be investigated. Thus there is the problem of spalling of the armor plate, or the shattering of the projectile. With somewhat different set of assumptions there is the problem of the dynamical enlargement of the circular hole, the major displacements of which take place plastically.

The problem to be discussed here represents still another approach which has the advantage of simplicity. It is essentially a quasi-dynamical analysis of the circular hole enlargement theory in that the displacement configuration under statical conditions is utilized. Such a procedure is frequently used in dynamical problems, for example, Rayleigh's method of calculating the fundamental frequency of oscillation of an elastic structure, or the first trial of an iteration procedure.

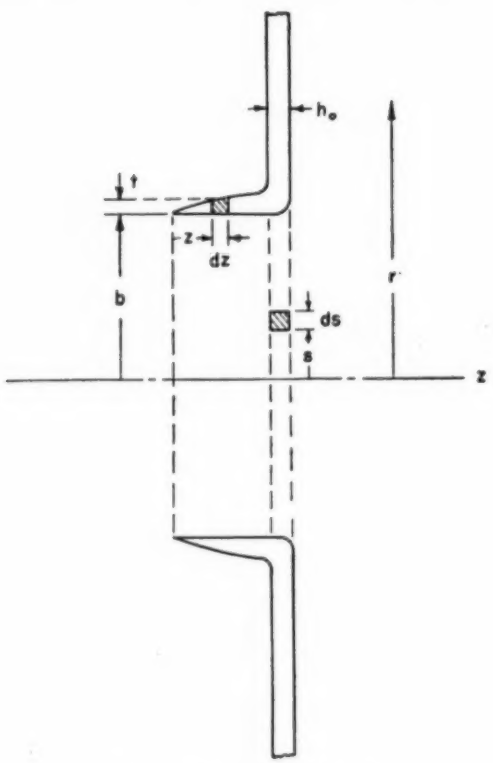


FIG. 1. Unsymmetrical mode.

We will consider here the two-dimensional problem where the thickness of the plate is small compared to the size of the hole. We will also assume that the projectile penetrates the plate without shattering. From intuitive considerations one might expect the material in the plate to form a crater, much like that formed when a nail is punched through a tin can.

We will here describe an experiment of G. I. Taylor[1] where a mandrel of gradual taper is forced through a thin plastic plate. He shows that two types of deformations can be obtained, the symmetrical and the unsymmetrical modes. Investigating the energy requirements for the two types of deformation, he finds that the energy of the symmetrical mode is 2.66 times that of the unsymmetrical mode, thereby inferring that the latter type of deformation is the more likely to be obtained.

Referring to Fig. 1, which shows the unsymmetrical mode, the following assumptions are made:

(1) Circumferential stress σ_θ is the only stress of significance in the crater so that σ_r and σ_z can be assumed to be zero.

(2) Plastic deformation then takes place under stress $\sigma_\theta = Y$, where Y is the yield stress.

(3) Plastic deformation takes place without a change of volume.

The first assumption of uniaxial stress requires that the strains ϵ_r and ϵ_z, perpendicular to σ_θ, be equal

$$\epsilon_r = \epsilon_z = \epsilon. \tag{1}$$

Thus an element $h_0 ds$ in the original position changes dimensions after deformation to the following:

$$t = h_0(1-\epsilon) \tag{2}$$

$$dz = ds(1-\epsilon). \tag{3}$$

From Eqs. (2) and (3) we have

$$ds/dz = h_0/t. \tag{4}$$

We next make use of the third assumption of no change in volume. By considering an elementary ring at s which is deformed to the crater,

$$2\pi s h_0 ds = 2\pi b t dz. \tag{5}$$

* Investigation supported by Rand Corporation, Santa Monica, California.

† Professor of Engineering, University of California, Los Angeles, California.

[1] G. I. Taylor, Quart. J. Mech. and Appl. Math. **1**, 103–124 (1948).

Substituting from Eq. (4), we have

$$t=h_0(s/b)^{\frac{1}{2}}. \tag{6}$$

Since $s/b=z/D$, the equation for the thickness at the crater becomes

$$t=h_0(z/D)^{\frac{1}{2}}. \tag{7}$$

To find D, the depth of the crater, we note that all of the material within the radius b must go into the crater. By equating these volumes,

$$2\pi b\int_0^D t\,dz=\pi b^2h_0. \tag{8}$$

Substituting for t from Eq. (7) and solving for D, we find

$$D=\tfrac{3}{4}b. \tag{9}$$

These are the results obtained by G. I. Taylor. Up to this point we have considered only deformations. We will next determine the work required for this deformation, including the dynamical effect.

Consider first the work required for the plastic deformation. Since the stress is uniaxial and equal to $\sigma_\theta=Y$, the work required per unit volume for deformation from s to b is

$$Y\int d\epsilon_\theta=Y\int_s^b\frac{ds}{s}=Y\ln\frac{b}{s}. \tag{10}$$

The total work done for the elementary ring is then

$$2\pi h_0 s\,ds\,Y\ln(b/s). \tag{11}$$

If the work done on all such rings from $s=0$ to $s=b$ is summed, the work required for the plastic deformation becomes

$$2\pi h_0 Y\int_0^b s\ln\frac{b}{s}ds=\tfrac{1}{2}\pi b^2h_0Y. \tag{12}$$

We next consider the dynamical effect. Let M be the total mass displaced at time t. Then the required accelerating force and the dynamical work done are

$$F=M\frac{d^2b}{dt^2}+\frac{dM}{dt}\frac{db}{dt} \tag{13}$$

and

$$W_d=\int M\frac{d^2b}{dt^2}db+\int\frac{dM}{dt}\frac{db}{dt}db. \tag{14}$$

Since M and dM/dt are

$$M=\pi\rho h_0b^2 \tag{15}$$

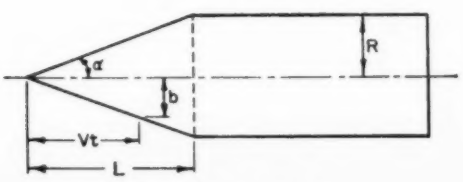


FIG. 2. Conical head.

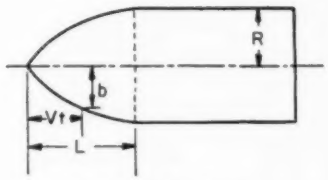


FIG. 3. Ogival head.

and

$$dM/dt=2\pi\rho h_0b(db/dt), \tag{16}$$

Eq. (14) becomes

$$W_d=\pi\rho h_0\int_0^{t_0}b^2\frac{d^2b}{dt^2}db+2\pi\rho h_0\int_0^{t_0}b\left(\frac{db}{dt}\right)^2db. \tag{17}$$

For the evaluation of Eq. (17) it is necessary to know the shape of the projectile head. We will consider here two shapes, the conical head and the ogival head.

(a) Conical Head

Figure 2 describes the conical head. The equations for b are $b=RVt/L$, $db/dt=RV/L$, and $d^2b/dt^2=0$. Substituting into Eq. (17) we obtain

$$W_d=2\pi\rho h_0(RV/L)^4(t_0^2/2)=\pi R^2h_0[\rho(VR/L)^2]. \tag{18}$$

Thus the total work done by a projectile of conical head becomes

$$W=\pi R^2h_0[\tfrac{1}{2}Y+\rho(VR/L)^2]. \tag{19}$$

(b) Ogival Head

The ogival head, shown in Fig. 3 can be described by the equations

$$b=R\sin\frac{\pi}{2}\frac{Vt}{L},$$

$$\frac{db}{dt}=\frac{\pi}{2}\frac{RV}{L}\cos\frac{\pi}{2}\frac{Vt}{L}$$

and

$$\frac{d^2b}{dt^2}=-R\left(\frac{\pi V}{2L}\right)^2\sin\frac{\pi Vt}{2L}.$$

The integrals in Eq. (17) are

$$\int_0^{t_0=L/R}b^2\frac{d^2b}{dt^2}db=\frac{\pi^2}{16}R^2\left(\frac{VR}{L}\right)^2$$

and

$$\int_0^{t_0=L/R}b\left(\frac{db}{dt}\right)^2db=\frac{\pi^2}{16}R^2\left(\frac{VR}{L}\right)^2.$$

Thus for the ogival head the total work required for penetration is

$$W=\pi R^2h_0[\tfrac{1}{2}Y+1.86\rho(VR/L)^2]. \tag{20}$$

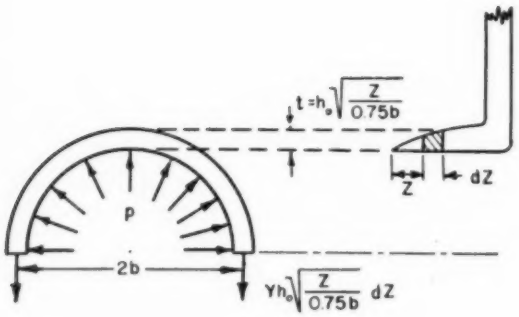


FIG. 4.

THE HEATING EFFECT

Because of friction between the projectile and the inner surface of the crater, heat will be generated. For its determination the pressure at z due to stress is found to be

$$p=\frac{Yh_0}{2b}\left(\frac{z}{0.75b}\right)^{\frac{1}{2}}. \tag{21}$$

In addition to this we must add the dynamical pressure which at z is

$$\rho h_0\left(\frac{z}{0.75b}\right)^{\frac{1}{2}}\frac{d^2b}{dt^2}. \tag{22}$$

There is one other term due to the rate of increase of the mass. This term can be assumed to act at the plane of the plate and is equal to

$$\rho(db/dt)^2. \tag{23}$$

If we let μ be the coefficient of friction between the sliding surfaces, the work dissipated by the friction force per unit time becomes

$$2\pi bV\mu\int_0^{0.75b}\left[\frac{Yh_0}{2b}\left(\frac{z}{0.75b}\right)^{\frac{1}{2}}+\rho h_0\left(\frac{z}{0.75b}\right)^{\frac{1}{2}}\frac{d^2b}{dt^2}\right]dz$$
$$+2\pi bV\mu\rho\left(\frac{db}{dt}\right)^2 h_0$$
$$=2\pi\mu h_0V\left\{\tfrac{1}{4}Yb+\rho\left[\tfrac{3}{4}b^2\frac{d^2b}{dt^2}+b\left(\frac{db}{dt}\right)^2\right]\right\}. \tag{24}$$

The amount of work dissipated in enlarging the hole from $b=0$ to b is then

$$W_q=2\pi\mu h_0V\left\{\tfrac{1}{4}Y\int_0^t bdt\right.$$
$$\left.+\rho\int_0^t\left[\tfrac{3}{4}b^2\frac{d^2b}{dt^2}+b\left(\frac{db}{dt}\right)^2\right]dt\right\}. \tag{25}$$

For evaluation it is necessary to know the shape of the projectile head. The quantity of heat equivalent to this work is

$$q=\frac{W_q}{9330}Btu, \tag{26}$$

and the temperature rise of the steel adjacent to the sliding surface becomes

$$\Delta T=\frac{q}{\rho C_p\Delta V}=\frac{W_q}{285\Delta V}°F, \tag{27}$$

where ΔV is the volume of material considered.

Because of the short duration of a ballistic penetration, the conduction of heat takes place only in a very thin layer of the projectile target interface. As a result the interface melts, and the molten metal serves as a friction-reducing lubricant which reduces W_q to a value just sufficient to maintain a melting temperature at the interface. Thus Eqs. (25) and (27) should be used with the reduced coefficient of friction due to the molten metal.

EXAMPLE

For the conical head, W_q at $t_0=L/V$ is

$$W_q=2\pi\mu h_0V\{\tfrac{1}{8}Y(RL/V)+\rho\tfrac{3}{8}(R^3V/L)\}.$$

The thickness h of the material which will reach or exceed a specified temperature rise ΔT is then found from Eq. (27) to be

$$(2\pi R^3_4R)h=W_q/285\Delta T.$$

Using the following numerical values: $\mu=.02$, $h_0=1.0$ in., $V=7000\times12$ in./sec, $Y=100\,000$ lb/in.^{2}, $R=3$ in., $L=6$ in., and $\rho=0.283/386=0.000733$ lb sec^2/in.^{4}, we find $W_q=1.38\times10^6$ lb in. and h in.$=1.38\times10^6/285\times42.4\Delta T=108/\Delta T$.

If we assume $\Delta T=2600°$F necessary for melting, the thickness of the melted interface becomes 0.042 in.

The work necessary for the irreversible plastic deformation of the target material is from Eq. (19) $W=38.0\times10^6$ lb in. With μ dropping to a low value of molten metal, the above figures indicate that the greatest portion of the energy is dissipated in the plastic deformation.

CONCLUSIONS

A quasi-dynamical approach of armor penetration leads to expressions for the energy dissipation for plastic deformation which are similar to those of the more rigorous analysis of the circular hole enlargement theory. Equation (20) compared to Eq. (19) indicates that the resistance to penetration for the ogival head is greater than that for the conical head. Heating of the projectile target interface was considered from dynamical pressures and assumed coefficient of friction. Because of melting at the interface the coefficient of friction must be greatly reduced after initiation of the melting action. A numerical example indicates that the energy dissipated by heating of the interface is generally small compared to the energy dissipated by the irreversible plastic deformation. In addition, energy would be dissipated by shock waves throughout the target plate.

JOURNAL OF APPLIED PHYSICS VOLUME 26, NUMBER 1 JANUARY, 1955

Spectroscopic Studies of Highly Ionized Argon Produced by Shock Waves*

H. E. Petschek, Peter H. Rose,† Herbert S. Glick,† Anne Kane, and Arthur Kantrowitz
Cornell University, Ithaca, New York

(Received December 31, 1953)

For the study of high-temperature gas dynamics, shock-tube techniques have been developed earlier to produce shock waves strong enough to heat gases to high and accurately known enthalpy (for argon up to 18 000°K or 40 percent ionization at equilibrium). This paper reports a study of the visible radiation from the highly luminous argon following strong shock waves.

Preliminary spectrograms showed a strong continuum and that the prominent argon lines were broadened and shifted to the red. Correlation of the frequency shifts with theoretical treatments permitted an evaluation of the ion density in the gas. Development of a drum camera spectrograph (film speed 700 ft/sec) in which time effects could be resolved to about 1 μsec indicated that equilibrium ion density was reached rapidly, and provided a rough measurement of the rate of decline of ion density due to cooling.

Absolute photoelectric spectrophotometric measurements of the continuum radiation were made and correlated with theoretical expectations. Confirmation of the expected variation of continuum intensity with wavelength and temperature was obtained and an undetermined factor on the theoretical intensity was evaluated. Determinations of the cooling rate of the high-temperature argon from time variation of the continuum intensity, the line shift, and the electrical conductivity (by others) are in agreement and show that continuum radiation was the dominant heat loss.

INTRODUCTION

THE dynamics of gases, with energy high enough so that appreciable dissociation and ionization can occur, has become important in several fields. In the flight of rockets, flow kinetic energies of this order (Mach No. ~10) have been approached. Astrophysicists have recently been devoting considerable attention to the violent dynamics of cosmical gas masses. Another example is provided by the gas dynamics of intense explosions and impulsive electrical discharges which produce shock waves strong enough to partially dissociate and ionize the gas.

High-temperature gas dynamics is also of interest because it can supplement the information on high-temperature gases from electrical discharge studies. For example, it is possible to prepare a high-temperature gas sample of accurately known enthalpy by using strong shock waves.

At high temperatures dissociation and ionization account for the major portion of the internal energy of the gas. Therefore the rates of approach to equilibrium of these processes will be important in the thermodynamics of the flow. Radiation will also be important in the heat balance and will sometimes dominate heat transfer problems. The appearance of an appreciable electron density can result in important long-range forces such as those leading to the phenomena of magneto-hydrodynamics and plasma oscillations. The long-range forces can produce radical departures from ordinary gas dynamics such as those seen in the solar corona.

This series of papers reports a study of some ionization phenomena accompanying strong shock waves in argon. A monatomic gas was chosen to eliminate the complication and energy loss of dissociation. Among the monatomic gases argon is most convenient because of its high molecular weight, low cost, and the large amount of atomic data available.

In a previous paper[1] an account was given of techniques which have been developed at Cornell for the production of strong shock waves in shock tubes. When the diaphragm between the high- and low-pressure section breaks, the expanding high-pressure gas acts like a piston pushing the low-pressure gas ahead of it. This creates a shock wave travelling at 4/3 to 9/8 (strong shocks in argon) the piston velocity. All the gas between the gas interface and the shock wave has at some time passed through the shock where its enthalpy was increased to $H \approx \frac{1}{2}u_1^2$ (Eq. 4, reference 1), where u_1 is the shock velocity. In coordinates travelling with the gas behind the incident shock, the closed end of the tube may be considered as a moving piston, which creates a reflected shock wave that moves away from the closed end. Across the reflected shock the enthalpy is increased by the flow kinetic energy of the gas behind the incident shock, i.e., the enthalpy is slightly more than doubled. Shocks were produced in argon travelling up to 17 times the speed of sound (Mach No. $M=17$) in the undisturbed gas. It was found that the shock speeds obtained agreed well with expectations from conventional one-dimensional shock-tube theory. Using the laws of conservation of mass, momentum, and energy, the state of the gas behind the shock can be calculated from the known state ahead of the shock and the measured shock velocity *if thermal equilibrium is assumed behind the shock*. Thus, equilibrium temperatures up to 18 000°K

* This work was largely supported by the U. S. Office of Naval Research.

† Many of the experiments reported here were performed by Mr. Rose, now at Douglas Aircraft Company, Santa Monica, California and Mr. Glick, now at Cornell Aeronautical Laboratory, Buffalo, New York, in partial fulfillment of requirements for the degree Master of Aeronautical Engineering.

[1] Resler, Lin, and Kantrowitz, J. Appl. Phys. 23, 1390 (1952).

and degrees of ionization up to 40 percent in argon are attainable.

It was found that a great deal of light was produced by strong shocks especially in argon so that spectroscopic study of the gas produced was possible. This study was undertaken to obtain information from spectra on the state of the argon following a strong shock wave. In particular, it was desired to determine how closely the gas approached the thermal equilibrium conditions assumed in the calculations given in reference 1. As the work progressed it was found possible to make time resolved spectra and absolute spectrophotometric measurements, so that the rate processes in the high-temperature gas could also be studied.

EXPERIMENTAL ARRANGEMENTS

The techniques of reference 1 were used for the production of strong shock waves in the argon. The high-pressure section was filled with H_2 or with H_2 burned with about 16 percent O_2 for greater shock strength. N_2 was added as a diluent to provide controlled reduction in shock strength. The low-pressure section of the shock tube was filled with commercial welding grade argon, 99.9 percent pure. During the course of this work, three separate shock tubes and spectrographs were used and the resulting spectra are referred to as first, second, and third series spectra. Spectrophotometric measurements were also made in conjunction with the third series. The shock tube and spectrograph arrangements are shown schematically in Fig. 1.

First Series

In the first-series spectra, the spectrograph looked along the axis of the shock tube. This had two advantages in an exploratory program. First, it was possible to use light from all parts of the heated gas between the shock wave and the contact discontinuity. Thus effects in all of these regions could be seen superposed. This superposition, however, created difficulties in quantitative analysis. Secondly, the greater optical path permitted a larger spectrograph exposure which was advantageous for the lower shock strengths.

The shock strengths in this series were from $M=5$ to $M=9$ as measured by the time to traverse the distance between two spark plugs.[1] Pressures ranging from 0.4 to 4 cm Hg were used. From the calculations of reference 1 (based on thermal equilibrium) this Mach number range corresponds to a temperature range from 5700°K to 12 000°K after the reflected shock. A small Constant Deviation Hilger Prism Spectroscope with an $f/20$ camera was employed in these experiments. The spectra were taken using Kodak Super Panchro Press Type B

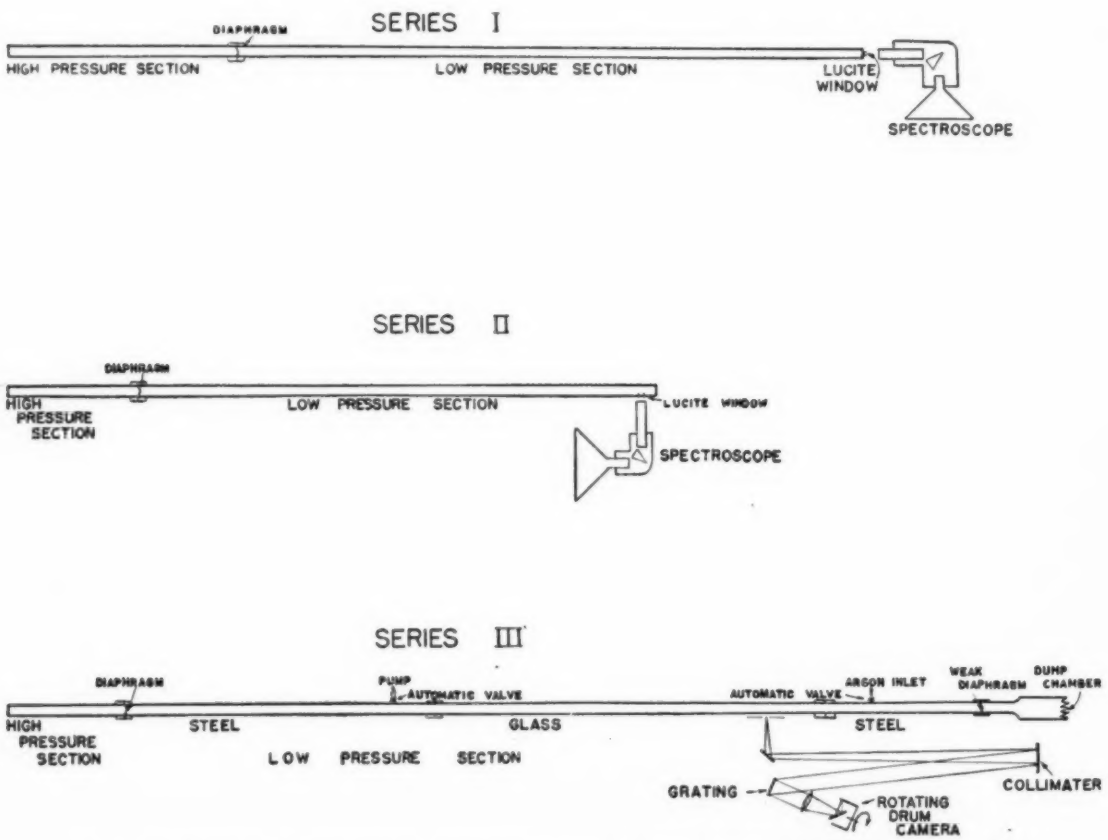


FIG. 1. Scale drawing of experimental arrangements for the three spectral series. Shock tube i.d. = 1½ in.

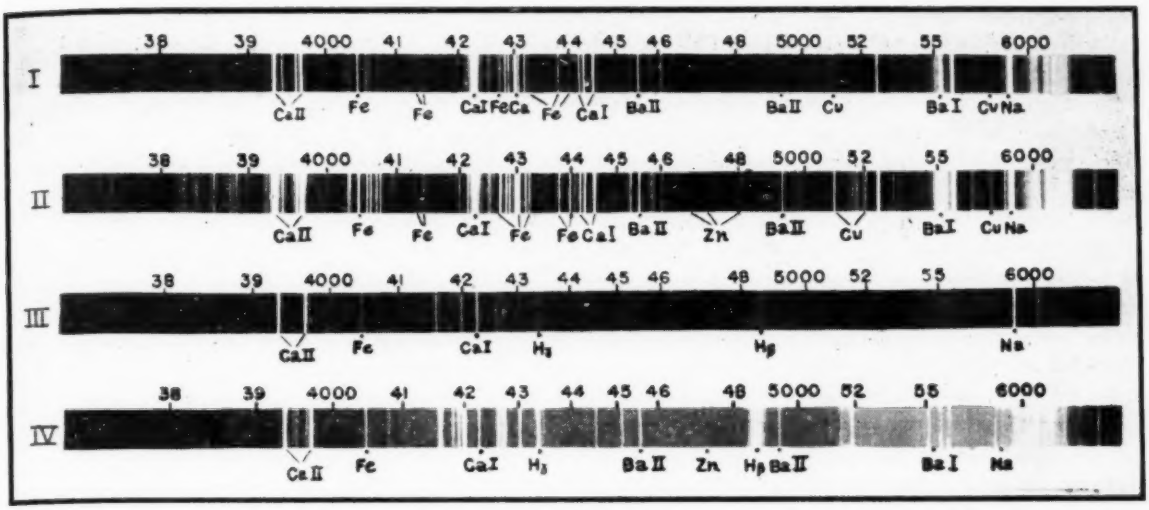

I
II
III
IV
38 39 4000 41 42 43 44 45 46 48 5000 52 55 6000
CaII Fe Fe CaI FeCa Fe CaI BaII BaII Cu BaI CuNa
CaII Fe Fe CaI Fe FeCaI BaII Zn BaII Cu BaI CuNa
CaII Fe CaI Hβ Hβ Na
CaII Fe CaI Hβ BaII Zn Hβ BaII BaI Na

FIG. 2.—Plate A. First series spectra.

Maximum theoretical temperatures and pressures observed by spectroscope.[a]

	M	P_1 in cm Hg	T_M in degrees K	P_M in cm Hg
I	5.34	4.03	6500	800
II	5.87	2.47	7800	575
III	7.22	0.764	10 200	270
IV	8.96	0.383	11 600	220

[a] Theoretical degree of ionization unknown due to high impurity level.

These spectra were taken using a brass shock tube in which high purity was not maintained. Note the increase with temperature of the intensity of lines of ionized forms, denoted by II, of Ca and Ba compared to that of lines due to unionized forms, denoted by I. Note also the broadening of H_β in spectrum III and its increase with temperature (see spectrum IV).

and Ansco Superpan Press films. A resolving power of about two angstroms at 4000 A and fourteen angstroms at 6000 A was obtained.

These early spectrograms [Plate A (Fig. 2)] were completely dominated by impurity emission, thus clearly demonstrating the necessity for considerable effort to improve the purity of the gas used in the low pressure region.

Second Series

In the second series of spectra, the purity of the working medium was improved in two ways. The shock tube was first evacuated to lower pressures, less than 0.01 cm Hg, and second, the general working pressure was raised (1 to 15 cm Hg). An interval of several minutes usually elapsed before the diaphragm burst so that the final impurity pressure is estimated at about 0.02 cm Hg. The shock strengths in these experiments ranged from $M=5.0$ to $M=12.5$ (temperatures after reflection 5700 to 17 000°K). The spectroscope used was similar to the one employed in the first series except that it had a denser glass prism. An $f/20$ camera using Kodak Panatomic X was employed, giving a resolving power of 0.4 A at 4000 A and 2 A at 6000 A. A reference argon Geissler-tube spectrum was superposed on a section of the shock spectrum for wavelength calibration. Some spectra with air as the low-pressure gas were also taken (Plate B, V).

In the second series spectra, argon lines broadened and shifted to the red were observed, and their quantitative interpretation as a measurement of the degree of ionization of the gas became an important objective. One of the difficulties in the first series was avoided by orienting the spectroscope to look across the shock tube (Fig. 1). In this case, only nonuniformities produced by boundary layers along the shock-tube walls are seen. The spectra also represent a time integration of the light emitted along the optical path. It would be expected, however, that except for self-absorption effects, the largest part of the spectrogram exposure would be attributable to the most highly excited gas visible to the spectrograph during the experiment. This gas is found following the shock reflected from the closed shock-tube end. Quantitative evaluation of the line shifts was possible from this series. It was clear, however, that more accurate results could be obtained if the necessity for unravelling the time-integrated effects could be avoided.

Third Series

To obtain time resolved spectra a grating spectrograph utilizing an air-driven drum camera was constructed.

The film speed used was about 700 ft/sec and the height of the slit (at the shock tube) was 0.25 in. This gave a possible resolving time of 1 microsecond (μsec). It should be noted that due to the gas velocity behind the shock wave particles that are seen at a time dt' apart in laboratory coordinates have passed through the shock wave at times dt apart where $dt=(u_1/u)dt'=(\rho/\rho_1)dt'$, where u_1 is the shock velocity, u is the particle velocity behind the shock relative to the shock velocity, ρ_1 is the density before the shock, and ρ is the density at the point being observed. Neglecting cooling of the hot gas, ρ/ρ_1 varies from 4 to 8 in the range of these experiments (see Fig. 3b in reference 1). Therefore, with a resolving time of 1 μsec, it is only possible to resolve times of 4 to 8 μsec in gas-particle history. To compromise between noise on the densitometer record and resolving time, a slit length on the densitometer equivalent to 3 μsec on the film (12 to 24 μsec in gas-particle history) was used. The noise level then permitted an accuracy of about ± 0.5 A in the wavelength of the intensity peaks of the shifted argon lines.

The grating used was a Bausch and Lomb original 30 000 lines per in. having an effective area of 4 square in. An $f/2.5$ aerial camera lens with a 7-in. focal length and Kodak spectroscopic film, type 103-F, were used. The resolving power of the spectrograms is estimated as 1 A over the entire spectrum. The reference spectra used in this series of experiments were argon and mercury. In order to obviate the possibility of a shift of the reference spectrum relative to the shock spectrum, the shock exposure was made in the middle of the reference exposure.

The third series spectra were taken with a shock tube modified in three ways (see Fig. 1). In the first place, a portion of the low-pressure region was replaced by a glass tube (Pyrex brand double tough pipe $1\frac{1}{2}$ in. i.d.) which reduced the effect of metallic dust particles. Secondly, a dump chamber was provided at the end of the shock tube to prevent the occurrence of reflected shock waves which might have destroyed the glass tube. The shock strengths in this series ranged from $M=11.2$ to $M=16.8$ (temperatures following the incident shock from 10 300 to 13 700°K). Thirdly, automatic valves permitted the use of flowing argon until the instant of diaphragm breakage thus considerably reducing the impurity level. The low-pressure section was first evacuated to pressures less than 0.01 cm Hg and then filled with flowing argon at a pressure of 1 cm Hg. The impurity level in this series, is, therefore, less than 1 percent. The impurity level may be appreciably lower due to details in the vacuum system design. In this series, all of the spectral lines observed were attributable to argon thus indicating a high gas purity.

Spectrophotometry

Spectrophotometric measurements were made simultaneously with the third series spectra. The spectroscope used was the same as the one used in the second series spectra and was oriented to look across the shock tube. A narrow spectral region was defined by slits of equal width at the collimator and focal plane. The resultant sensitivity profile is an isoceles triangle with a base of about 6 A at 4200 A. The light intensity was measured by a GL 929 photocell and recorded on a Tektronix 512 oscilloscope. For weak shock waves, a preamplifier with a gain of ten was necessary.

The rise time of the electronic equipment was measured from a short duration spark and was ~ 2 μsec. The over-all rise time in the experimental setup was, however, not quite this good since the spectroscope collimator received light from about 0.2 in. along the shock tube. Thus a region of the tube is seen which corresponds to ~ 3 μsec in laboratory coordinates. The measured rise time on some high Mach number traces, where equilibrium is attained very rapidly, was ~ 5.7 μsec.

An absolute intensity calibration of the spectroscope and photocell was made using the sun as a standard. Abbot's data[2] for the intensity of the sun's radiation at the surface of the earth were used. An estimate of the calibration accuracy was made by comparing measurements made on three separate days. The most probable error is ± 20 percent.

GENERAL CHARACTERISTICS OF SPECTRA

In the first series spectra hydrogen lines H_β and H_γ were found considerably broadened and the half-widths increased with temperature [see Plate A, III and IV (Fig. 2)]. Following a suggestion of Dr. R. Ladenburg this broadening was attributed to Stark effect due to the electrical fields of nearby ions. It was found that the experimental half-widths were in rough agreement with the first-order Stark effect broadening[3] which would be expected under our experimental conditions. However, due to uncertainties in our knowledge of the gas state near the H_2-A interface and some doubts about the line-broadening theory (see Line Broadening and Shift section) this effect was not explored further. The hydrogen lines did not appear in the second series spectra indicating that the region emitting these lines is in the vicinity of the H_2-A interface, which was not visible to the spectroscope in the second series (see Fig. 1).

In the second series spectra our attention was directed primarily to the study of the prominent argon lines which were broadened and shifted to the red [see Plate B, I and II (Fig. 3)]. The measurement of the frequency shift was facilitated by superimposing an argon reference spectrum on the shock and making a densitometer record of the spectra. The measured frequency shifts (up to four A) were of the order of magnitude we would expect, assuming thermal equilibrium and agreed with the Baranger theory (see next section) over a limited temperature range. This range was limited on the low-

[2] C. G. Abbot, Annals of the Astrophysical Observatory of the Smithsonian Institute Vol. II–V (1908–1932).

[3] Similar results have been obtained by Fowler, Atkinson, and Marks, Phys. Rev. 87, 966 (1952).

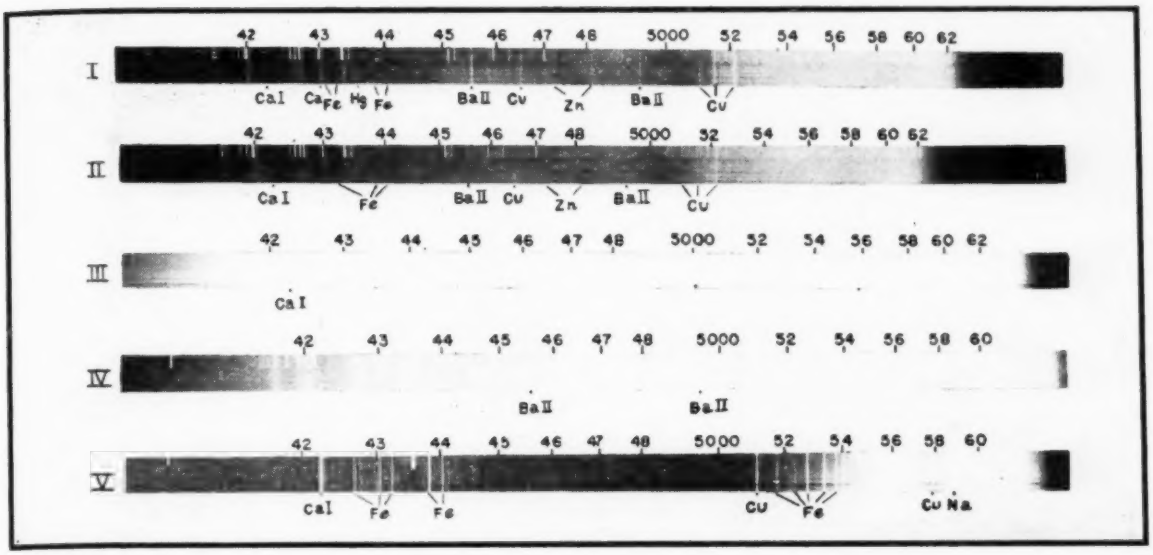

I
II
III
IV
V
42 43 44 45 46 47 48 5000 52 54 56 58 60 62
CaI Ca Fe Hg Fe BaII Cu Zn BaII Cu
CaI Fe BaII Cu Zn BaII Cu
CaI
BaII BaII
CaI Fe Fe Cu Fe Cu Na

FIG. 3.—Plate B. Second series spectra.

Maximum theoretical temperature, pressure, and degree of ionization observed by spectroscope.

	M	P_1 in cm Hg	T_M in degrees K	P_M in cm Hg	Degree of ionization	Reference spectrum used
I—Argon	7.30	5.00	11 000	1800	0.014	Argon
II—Argon	8.12	5.00	12 000	2250	0.025	Argon
III—Argon	12.52	5.00	16 700	6600	0.172	Argon
IV—Argon	8.47	5.00	12 500	2450	0.031	Shock tube argon
V—Air	9.55	5.00	6000	4000	0.30[a]	Mercury

[a] Degree of dissociation.

I, II, and III are a selection from the second series spectra in which intermediate purity was maintained to demonstrate the effects of increasing temperature on the shock spectrum of argon. The broadening and displacement of the argon lines can be seen in I and II in comparison with the superposed argon Geissler-tube spectrum. Note also the increase in the intensity of the continuous background with increasing temperature (I→II→IV→III). Spectrum IV shows a shock spectrum in fairly pure argon and a superposed spectrum from a Geissler tube filled with a sample of the shock tube gas. It will be seen that the argon spectrum was not excited in the Geissler tube and the prominent band spectra which were excited did not appear in the shock spectrum indicating complete dissociation. Spectrum V is an air spectrum showing the impurity line spectrum characteristic of low temperatures and prominent bands in the yellow and red.

temperature end by insufficient exposure and at the high-temperature end by effects of rapid cooling which made interpretation of the time integrated spectra very difficult. These results have since been superseded by more quantitative third series data.

Typical third series spectra are to be seen in Plate C (Fig. 4). It can be seen that the line broadening and shift reach a maximum very quickly behind the onset of luminosity.‡ At the lowest Mach numbers a suggestion of a noticeable ionization rise time could be seen on the film. Measurements of the maximum shifts are plotted in Fig. 5 and are discussed in the next section.

The cooling of the gas can be observed in Plate C (Fig. 4) from the decreasing broadening and shift of the argon with time. Measurements of the rate of decrease of ionization from the argon line shift have been made. The cooling of the gas also results in a rapid drop in the continuum intensity. More accurate measurements of this were made with the spectrophotometer. Quantitative cooling results are given in the Cooling Rate section.

In the neighborhood of 6000 A there is a broad band in the third series spectra. This is apparently due to the argon lines 6032, 6043, 6059 A. These lines all have much larger Stark coefficients than the lines around 4000 A, which accounts for the extreme broadening. Unfortunately, no quantitative work could be done with these lines since, in the experimental range covered, neither the Holtzmark nor the Baranger (see next section) theory applies.

The intensity ratio of certain prominent lines in the first and second series varied noticeably with the temperature. On Plate A (Fig. 2) compare Ba II 4934.1 and Ba II 4554.0 with Ba I 5535.6; Ca II 3933.7 and Ca II 3968.5 with Ca I 4226.7. The Ba II to Ba I and Ca II to Ca I intensity ratios increased with temperature as in the transition from arc to spark spectra. On Plate B (Fig. 3) compare the argon lines with the background. It will be seen that the intensity of the continuum rises relative to the line intensities.

In the second series, as the temperature increased, a

‡ Recent work (to be published later) has shown that there is a time lag between the shock wave and the onset of luminosity. Due to the sharpness of the rise of luminosity it was for some time mistakenly identified with the shock wave.

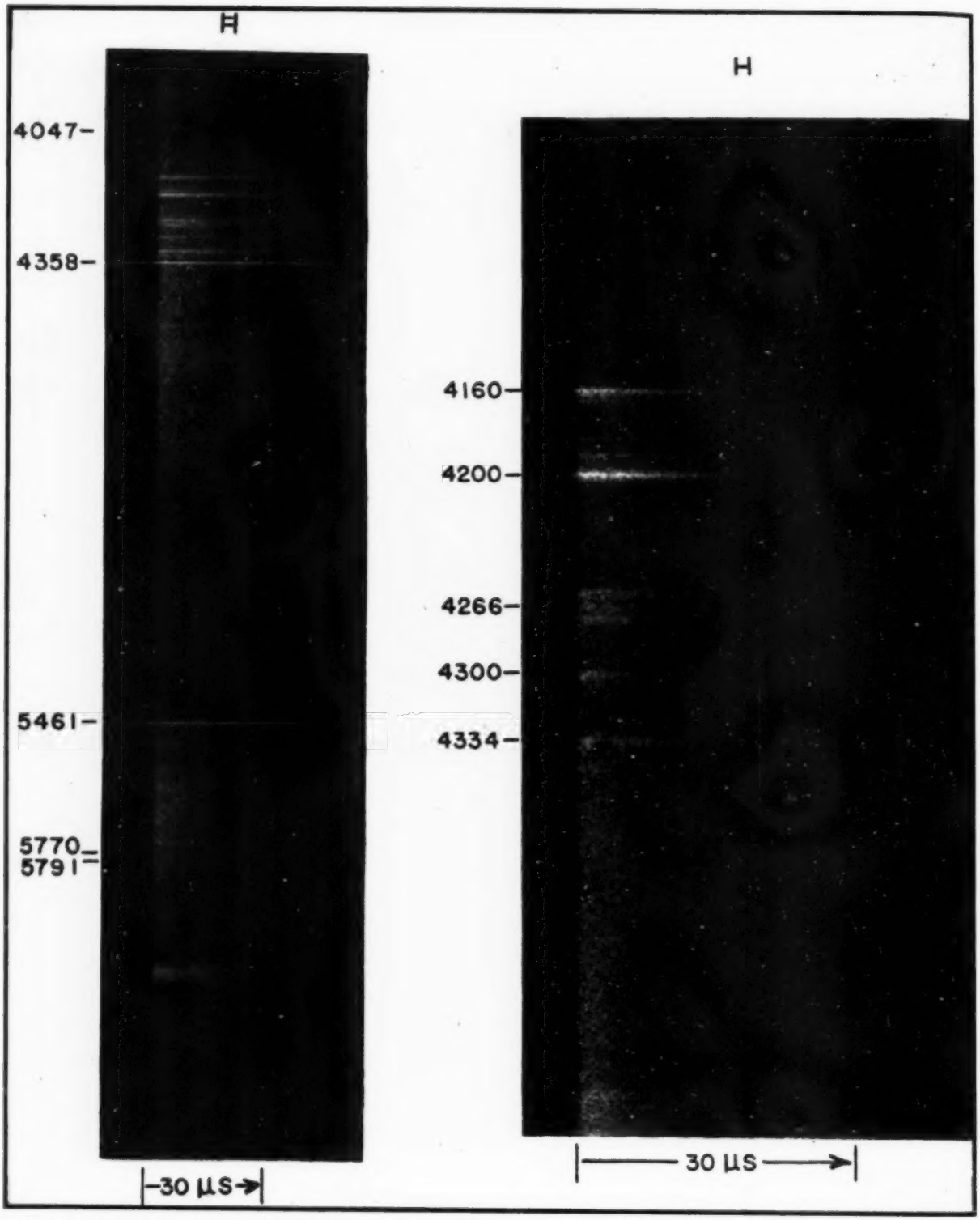


FIG. 4.—Plate C. Third series time resolved spectra.

Maximum theoretical temperature, pressure, and degree of ionization observed by spectroscope.

	M	P_1 in cm Hg	T_M in degrees K	P_M in cm Hg	Degrees of ionization	Reference spectrum used
I—Argon	15.9	1	13 300	360	0.137	Argon
II—Argon	13.9	1	12 200	270	0.081	Mercury

I is an enlargement of the region around 4200 A with a superposed argon Geissler-tube reference spectrum. Note that the line broadening and shift (to the red) is a maximum at the onset of luminosity and decreases with time as the gas cools. Note also the sharp reduction in continuum intensity with time. II shows the entire visible spectrum with a superposed mercury Geissler tube reference spectrum. The broadening decreases less rapidly with time than in I due to the lower temperature. Note also the very broad lines in the red (~6000 A). All the lines in the shock spectrum are argon lines.

few of the impurity lines reversed (notably the Na D lines, Ca I 4226.7, and C_2 5165.2, etc.)§ and others blurred into the background until at 16 000°K only reversed lines were sharp. Mercury and zinc lines with higher excitation potentials than the other impurity lines appeared only at the higher temperatures. Ba I and Ti I (but not II) lines disappeared as the temperature increased indicating that these atoms become completely ionized.

No detailed studies of the intensity of impurity lines were made because the density of the impurity was unknown and likely to vary from run to run. The intensity of the argon lines, however, can give useful information about the population of excited states and should be studied further.

We ran a number of experiments with a Geissler-tube comparison by running the argon gas from the low-pressure end of the tube into a Geissler tube and superimposing this spectrum on the shock spectrum Plate B, IV (Fig. 3). The Geissler-tube spectrum showed mostly bands of molecules which were apparently more easily excited than argon and probably were completely dissociated in the shock.

A few shocks in the second series were run with air in the low-pressure end. Again, the spectra showed atomic lines of impurities to be prominent in addition to bands (or continua) in the green through red region. These unidentified bands were also observed in the first series "argon" spectra [Plate A, I and II (Fig. 2)].

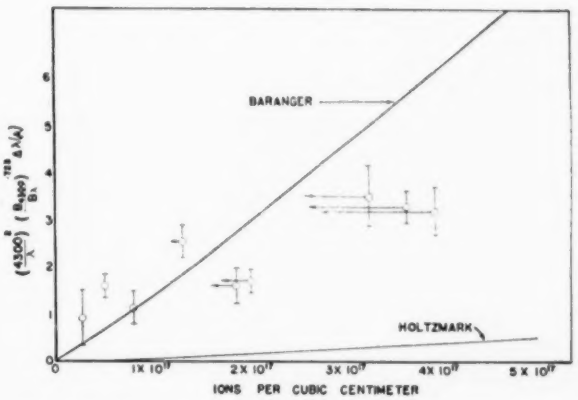


FIG. 5. A comparison of theoretical and experimental wavelength shifts from third series spectra [see Plate C (Fig. 4)]. The Holtzmark theoretical curve was obtained using measured Stark coefficients and the electric fields from a random static distribution of ions. The Baranger curve includes the effect of electron collisions which are more important in this range. The plotted positive ion density for the experimental points was obtained from the measured shock velocity and initial pressure (1 cm Hg) assuming thermal equilibrium. Wavelength shifts were measured for the following lines 4345.2, 4333.6, 4300.1, 4272.2, 4266.3, 4259.4, 4200.7, 4198.3, and 4191.0 A and were normalized to correspond to the 4300 A line by Eq. (1). The arrows indicate reduction in ion concentration due to radiation cooling (Fig. 9) during the measurement time (2 μsec). The final agreement indicates that equilibrium ionization was attained.

LINE BROADENING AND SHIFT

The broadening and shift of spectral lines due to perturbation of the energy levels by particles in the neighborhood of the radiating atom has been considered by several people. Lorentz,[4] Weisskopf,[5] Lindholm,[6] Foley,[7] and others, have calculated the broadening and shift under the assumption that the time of a collision is short compared to the time between collisions (low density and short-range forces). The effect of the collisions may then be considered as an instantaneous change in the phase of the radiation.

Another extreme applicable to high densities and long-range forces has been formulated by Holtzmark.[8] He calculates the distribution of the electric field at a point due to a random distribution of charges, dipoles, or quadrupoles. Then assuming that the field is static, the shape of the line can be found from the calculated or measured Stark coefficient. In order to consider the field static, it is necessary that the characteristic time τ_c of variation of the field be so long that the indeterminacy in frequency $1/\tau_c$ of a finite wave train of length τ_c is much less than the change in frequency ω due to the electric field ($\omega\tau_c \ll 1$).

The experimental conditions treated here, however, do not fit either the Lorentz or the Holtzmark conditions directly. Due to the high degree of ionization we need consider only interactions of charged particles. The thermal velocities of the ions are low enough so that their fields may be considered static ($\omega\tau_c \approx 10$ for the observed ω). For the electrons, however, $\omega\tau_c = 1/20$. Furthermore, for inverse square law forces it is not possible to assume that the time of collision is small compared to the time between collisions.

Baranger[9] has treated this case by dividing the effect of the electrons into two groups. He defines a critical radius surrounding the radiating atom such that electrons passing within this radius produce a sufficiently rapidly varying field to be considered instantaneous. This radius can be chosen (over our experimental range) such that there is only one electron within it at a time and such that the phase change caused by a single electron at the critical radius is negligible. The cumulative effect of all the electrons outside this radius is still appreciable but may be considered as static. The line contour can then be calculated by considering the effect of the close electron collisions on a spectral line in a static field and then summing over the distribution of the static field caused by the ions and distant electrons.

§ The C_2 bands only appeared in absorption and were probably produced by evaporation products from the Lucite window used.

[4] H. A. Lorentz, Proc. Acad. Sci. (Amsterdam) **8**, 591 (1906).
[5] V. Weisskopf, Z. Physik **80**, 423 (1933).
[6] E. Lindholm, Arkiv Mat., Astron. Fysik **28B**, No. 3 (1942).
[7] H. M. Foley, Phys. Rev. **69**, 616 (1946).
[8] J. Holtzmark, Ann. Physik **58**, 577 (1919).
[9] M. Baranger, Phys. Rev. (to be published). Since this theory is as yet unpublished, it should be pointed out that Foley's theory gives displacements only 20 percent less than Baranger's. Slightly more accurate measurements than ours would be required to differentiate between these theories.

Since in a collision the direction as well as the magnitude of the electric field changes with time, the calculation is considerably simplified for lines whose Stark coefficient is independent of the magnetic quantum number. This condition is satisfied by a considerable number of visible argon lines. A further restriction imposed by the theory is that the probability of inducing a transition to a neighboring state by an electron collision is small. A list of the observed argon lines to which this theory is applicable is given in the title to Fig. 5. The calculated displacement of the peak intensity is

$$\Delta\lambda = 2\times 10^{-15}\lambda^2 N_e^{\frac{1}{4}} T^{\frac{1}{4}} \left(54\pi\times 10^{-7}\frac{bN_e}{\sqrt{T}}\right)^{0.723}, \quad (1)$$

where λ is the wavelength in A, N_e the electron density in cm^{-3}, T the absolute temperature in °K, and b the Stark coefficient in $cm^{-1}/(volt/cm)^2$.

In Fig. 5 the theoretical shift of the 4300 A line has been plotted as a function of ion density for the equilibrium conditions behind a shock wave propagating in room-temperature argon at 1 cm Hg pressure.

The ion densities for the experimental points in Fig. 5 were obtained from the shock velocity as measured with (a second) drum camera, conservation of mass, energy, and momentum across the shock, the perfect gas equation of state, and the Saha equation for the degree of ionization, assuming thermal equilibrium. These calculations can be found in the appendix of reference 1. The wavelength shifts for the experimental points were found from densitometer records made from third series spectra with a surveying slit length equivalent to 3 μsec on the film. The densitometer record was made at the region of maximum ion density. The measured shifts in wavelength have been multiplied by the theoretical ratio [from Eq. (1)] of the shift of the 4300 A line to the shift of the line being measured so that a single plot will include all the lines measured. The indicated error is the standard deviation from the mean of the results from the various lines measured from a single spectrogram.

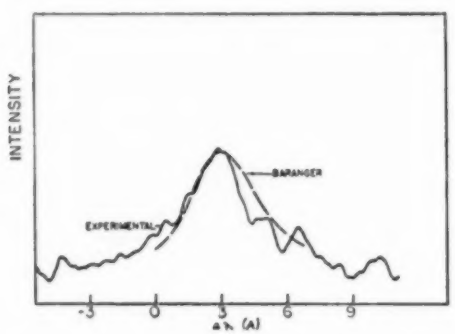


FIG. 6. Comparison of theoretical and experimental 4300 A line contours for $T\approx 11\,500$°K. The experimental contour is obtained from a densitometer record without correction for film or densitometer nonlinearity. For more convenient contour comparison the theoretical peak intensity and wavelength shift were adjusted to equal the experimental values.

The Stark coefficients for argon employed were measured by Minnhagen.[10]

At high Mach numbers it was found necessary to make a correction due to the cooling of the gas within the resolving time of the equipment. The correction has been made from calculations similar to those used in constructing Fig. 9, assuming 2 μsec cooling time (8–16 μsec in gas particle history). The equilibrium theoretical ion density has been plotted as abscissa in Fig. 5 with an arrow indicating the calculated change in ion density due to cooling.

Figure 5 shows that the equilibrium degree of ionization is attained behind the shock wave in this experimental range.

In Fig. 6 the theoretical line shape is compared to a densitometer record of the 4300 A line on one of the third series spectra. The curves have been normalized to give the same shift and intensity at the peak. In view of the fact that linear response of the film and densitometer have been assumed in drawing Fig. 6 the agreement may be partially fortuitous.

CONTINUUM RADIATION

The theory of the continuum radiation from a partially ionized gas as formulated by Unsöld[11] has been employed here. He assumes a hydrogenlike atom and uses the Kramer's[12] absorption formula to calculate the absorption coefficient per atom with an equilibrium distribution of excited states (free and bound). Summation over the excited states is carried out with the assumption that the states are so close together that the summation may be replaced by an integration. Assuming equilibrium the absorption is converted into the emission.

It was necessary for Unsöld to include two rather uncertain factors in this theory. The first is due to the fact that in an ionized gas the energy levels near the ionization limit are smeared into one another. Thus transitions from these levels to lower ones give rise to continuum rather than line radiation. The energy range ΔE over which the levels are smeared together has been calculated for hydrogen and the alkalies by Inglis and Teller.[13] The result is, however, not directly applicable to argon. Secondly, in order to apply the theory to atoms other than hydrogen it is necessary to assume an effective nuclear charge z_{eff}.

The final result for the radiation per cubic centimeter per unit frequency interval is

$$I_\nu = \gamma\frac{128\pi^3}{3\sqrt{3}}\left(\frac{e^2}{hc}\right)^3 z_{\text{eff}}^2 \times\exp(\Delta E/kT)N_A kT\exp(-T_{\text{ion}}/T), \quad (2)$$

[10] L. Minnhagen, Arkiv Fysik **1**, No. 20, 425 (1949).
[11] A. Unsöld, Ann. Physik **33**, 607 (1938).
[12] H. A. Kramers, Phil. Mag. **46**, 836 (1923); J. A. Gaunt, Proc. Roy. Soc. (London) **A126**, 654 (1930); A. W. Maue, Ann. Physik **13**, 161 (1932); D. H. Menzel and Ch. L. Pekeris, Monthly Notices Roy. Astron. Soc. **96**, 77 (1936).
[13] D. E. Inglis and R. Teller, Astrophys. J. **90**, 439 (1939).

where γ is the ratio of the internal partition function of the ion and electron to the internal partition function of the ground state of the atom. For argon $\gamma \approx 12$ in our temperature range. e, h, c, and k are the electronic charge, Planck's constant, the velocity of light, and Boltzmann's constant, N_A is the number density of neutral atoms, and T_{ion} is the ionization potential expressed in °K. This expression may be rewritten in a more convenient form by substitution of the Saha equation (neglecting excited electronic states)

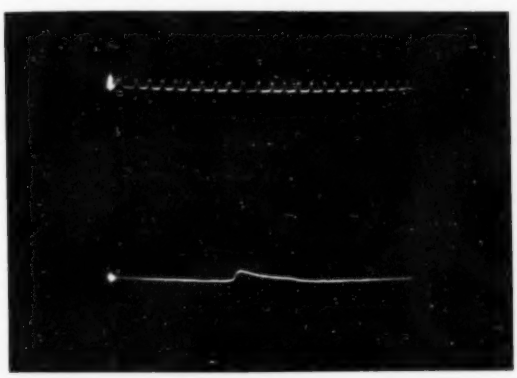
$$\frac{N_e^2}{N_A} = \frac{(2\pi mkT)^{\frac{3}{2}}}{h^3} \gamma \exp(-T_{\text{ion}}/T), \tag{3}$$

to obtain

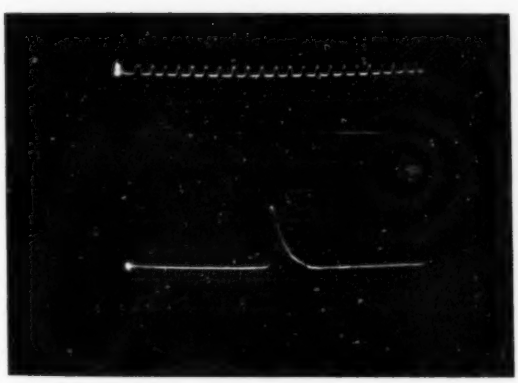
$$I_\nu = \frac{64\pi^{\frac{3}{2}}}{3\sqrt{6}} \frac{e^6}{m^{\frac{1}{2}}c^3} z_{\text{eff}}^2 \exp(\Delta E/kT) \frac{N_e^2}{(kT)^{\frac{1}{2}}}, \tag{4}$$

where m is the electron mass.

It will be noted that the intensity is independent of frequency. This result is not valid for frequencies which would correspond to transitions to levels where the energy difference between adjacent stationary states is large; i.e., where the replacement of the summation over states by an integral is not a good approximation. In argon the first excited state occurs at 11.5 ev or 4.2 ev below the ionization limit. Since the capture of electrons with energy greater than about $\frac{3}{2}kT$ to the first excited state will be fairly infrequent, we may expect that the continuum radiation will cut off well beyond the region accessible to us at about 6 ev, i.e., $\lambda = 2000$ A.[14]



(a)



(b)

FIG. 7. Spectrophotometer oscillograms. These are samples of the spectrophotometer oscillogram traces from which the continuum intensity at 5000 A° and the rate of cooling of the gas was measured. Note that in (a) ($M = 11.9$ $T = 10\,900$°K at peak) there is a noticeable rise time for the buildup of luminosity and the cooling is slow. In (b) ($M = 16.8$ $T = 12\,900$°K at peak) the rise time is limited by the experimental equipment and rapid cooling is to be seen. The timing marker shown is a 100 kc oscillator. The sensitivity in (b) is $\frac{1}{3}$ that in (a).

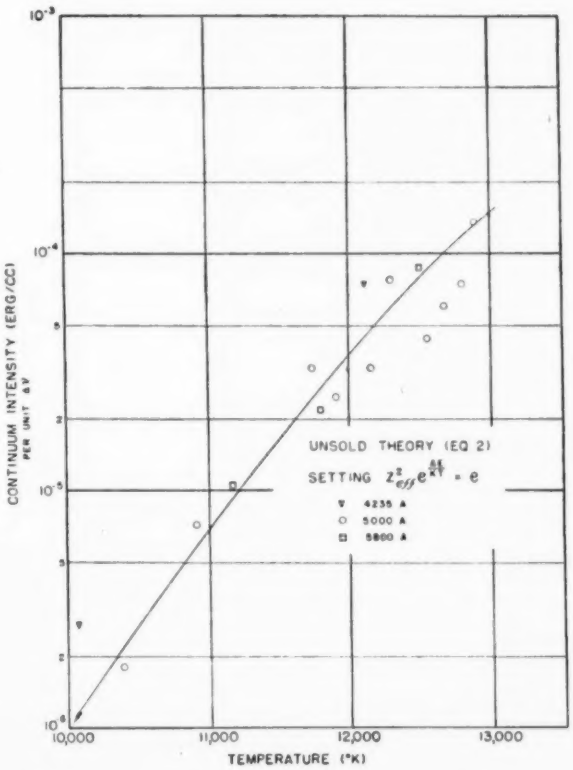


FIG. 8. Comparison between spectrophotometric measurements and theoretical continuum radiation after adjusting an undetermined factor. The independence of radiation intensity on frequency is in agreement with theoretical expectations. (The density is defined by $\rho_1 = 1$ cm Hg.)

Some examples of the oscillograms obtained with the spectrophotometer are shown in Fig. 7. Over most of the Mach number range the rise time of the continuum intensity signal was unfortunately limited by the rise time of the equipment. Near the bottom of the range $M = 11$ the rate of buildup of radiation in the gas became barely noticeable (see Fig. 7a). All of the traces show a decreasing intensity with time corresponding to the cooling of the gas.

From the (temporal) peaks of the continuum intensity traces it has been possible to verify the general characteristics of the Unsöld theory and to obtain a rough estimate of the undetermined product $z_{\text{eff}}^2 \exp(\Delta E/kT)$. This is based on the results of the third series spectra that equilibrium is attained in the range of these measurements, which makes a calculation of the state of the gas possible. In Fig. 8 the experimentally meas-

[14] Measurements of continuum radiation from a hydrogen arc have been made by G. Jürgens, Z. Physik 134, 21 (1952).

ured peak intensity is plotted as a function of temperature. It will be seen that the intensity at the three wavelengths, 4235 A, 5000 A, and 5800 A (where no lines were found in the spectra) is equal well within the scatter, in agreement with the Unsöld-Kramers theory. The theoretical curve has been plotted with $z_{eff}=1$ and $\Delta E/kT=1$. (More accurately $z_{eff}^2\exp(\Delta E/kT)=e$.) Good agreement is attained with this factor in the temperature range from 10 000 to 13 000°K after a correction based on Fig. 9 due to cooling of the gas within the resolving time of the spectrophotometer.

This cooling correction is derived from the fact that at the peak of the trace the value of the intensity appropriate to that time is recorded. For a sharply peaked input this maximum can occur in less than the rise time of the equipment. At high Mach numbers where the largest correction is necessary the peak occurs at $\sim$2.5 μsec after the onset of luminosity resulting in a temperature correction of 800°K. The theoretical curve was plotted with a density appropriate to 2.5 μsec after a shock wave propagating into argon at room temperature and 1 cm Hg pressure.

COOLING RATE

Several of the experiments which have been performed (spectrophotometry, electrical conductivity, and line shifts) show a decline in the degree of ionization with time due to cooling. In this section we shall utilize these observations to determine the heat loss and to compare it with the computed loss due to continuum radiation. To relate the change in the degree of ionization to the heat loss, it is necessary to take into account energy exchanges with the dynamics of the gas.

Since any waves propagating in the gas behind the shock would cause changes in the shock velocity, the flow may be considered as a steady flow in coordinates moving with the shock wave if the shock wave moves at a constant velocity. The drum camera pictures of the shock propagating down the tube show a slight curvature particularly at high Mach numbers. The largest rate of change of shock velocity measured was about 0.1 percent per centimeter. Since the resulting rate of change of degree of ionization is less than 10 percent of the rate of change due to cooling, we will assume that the nonsteady effects are negligible. Considering the flow behind the shock as a steady one-dimensional constant area channel flow we may write the equations for conservation of mass

$$\rho u=\rho_1 u_1, \tag{5}$$

and momentum

$$p+\rho u^2=p_1+\rho_1 u_1^2\approx\rho_1 u_1^2, \tag{6}$$

where the subscript 1 denotes conditions before the shock wave. The equation of state in an ionized gas is

$$p=(1+\alpha)\rho RT, \tag{7}$$

where α is the fraction of the original particles ionized, and R the gas constant. For $\alpha^2\ll 1$ the Saha equation may be written as

$$\alpha=A\frac{T^{5/4}}{\sqrt{p}}\exp\left(-\frac{T_{\text{ion}}}{2T}\right), \tag{8}$$

where A is a constant.

The differential forms of these equations may be written as

$$u d\rho=-\rho du, \tag{5'}$$

$$dp=-\rho_1 u_1 du, \tag{6'}$$

$$\frac{1}{p}dp=\frac{d\alpha}{1+\alpha}+\frac{d\rho}{\rho}+\frac{dT}{T}, \tag{7'}$$

$$\frac{d\alpha}{\alpha}=-\frac{1}{2p}dp+\left(\frac{5}{4}+\frac{T_{\text{ion}}}{2T}\right)\frac{dT}{T}. \tag{8'}$$

By algebraic manipulation of the integral and differential forms we obtain

$$d\alpha=\left[\frac{\dfrac{5}{4T}+\dfrac{T_{\text{ion}}}{2T^2}+\dfrac{1}{2T\left(\dfrac{u_1}{u}-2\right)}}{\dfrac{1}{\alpha}-\dfrac{1}{2(1+\alpha)\left(\dfrac{u_1}{u}-2\right)}}\right]dT=cdT, \tag{9}$$

and similar expressions relating the other differentials. The energy equation may be written as

$$\frac{d(H+\frac{1}{2}u^2)}{dt}=u\frac{d(H+\frac{1}{2}u^2)}{dx}=-\frac{Q}{\rho}, \tag{10}$$

where

$$H=\frac{5}{2}(1+\alpha)RT+\alpha RT_{\text{ion}}$$

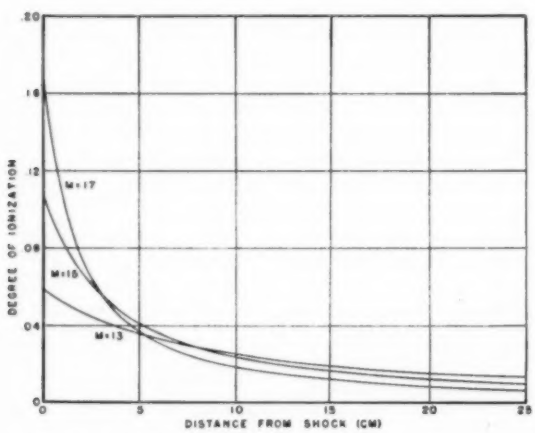


FIG. 9. Calculated variation of degree of ionization as a function of distance behind the shock wave due to continuum radiation cooling ($\rho_1=1$ cm Hg). These calculations were employed to make allowances for the finite resolving time in these experiments.

is the enthalpy per unit mass, t is the time, x the distance from the shock wave, and Q is the rate of energy loss per unit volume. Performing the differentiation in Eq. (10) and substituting the relations of the type in Eq. (9) we obtain

$$\frac{d\alpha}{dx}=-\frac{Q/\rho}{uR\left[\left(\mathrm{T}+\frac{1+\alpha}{c}\right)\frac{\frac{5}{2}-\frac{4}{u_1}}{1-\frac{2u}{u_1}}+T_{\text{ion}}\right]}, \tag{11}$$

and similar expressions for the rate of change of the other variables. Given Q, the state of the gas as a function of distance behind the shock may be calculated by a step-wise integration.

If the cooling is due to continuum radiation, Q is given by

$$Q=I_\nu\Delta\nu, \tag{12}$$

where $\Delta\nu$ is the cut-off frequency mentioned in the last section ($h\Delta\nu=6$ ev) and I_ν is given by Eq. (4) with $z_{ef}^2\exp(\Delta E/kT)=e$. The energy loss due to capture to the ground state will be negligible under our conditions due to the blackbody limit, i.e., self-absorption. In Fig. 9, the degree of ionization is shown as a function of distance behind the shock wave for three different Mach numbers. The changes in all quantities due to cooling increase as the shock velocity increases. At $M=17$ the pressure increases by 6 percent in 25 cm. The temperature at the three Mach numbers decreases to about 10 800°K in 10 cm and to 10 000°K in 25 cm. These calculations were made taking steps of $\Delta\alpha/\alpha=-0.15$.

In order to use our experimental data as measurements of the heat loss it is convenient to transform Eq. (11) so that the characteristic cooling time, defined as the energy invested in ionization divided by rate of energy loss, is

$$\frac{N_e kT_{\text{ion}}}{Q}=-\frac{\rho}{\rho_1}\left[\frac{N_e k\left(T+\frac{1+\alpha}{c}\right)}{\rho_1 u_1^2\left(1-\frac{2u}{u_1}\right)}\cdot\frac{1-\frac{\frac{5}{2}-\frac{4u}{u_1}}{1-\frac{2u}{u_1}}}{1+\left(\frac{T}{T_{\text{ion}}}+\frac{1+\alpha}{cT_{\text{ion}}}\right)\frac{\frac{5}{2}-\frac{4u}{u_1}}{1-\frac{2u}{u_1}}}\right]\frac{1}{\frac{1}{N_e}\frac{dN_e}{dt'}}, \tag{13}$$

where t' is the time at a fixed point in the laboratory since the passage of the shock (see Experimental Arrangements section). Since the term in the parenthesis only varies from 0.63 to 0.44 in our experimental range and ρ only varies slightly as the gas cools, the characteristic cooling time is essentially determined by $1/N_e\,dN_e/dt'$.

Since it was shown in the last section that the continuum intensity is proportional to the ion density squared, (Eq. (4))

$$\frac{1}{I_\nu}\frac{dI_\nu}{dt'}=\frac{2}{N_e}\frac{dN_e}{dt'}, \tag{14}$$

if the variation of $1/\sqrt{T}$ is neglected. The quantity on the left can be measured from the spectrophotometer oscillograms. The deflection, slope, and time since the passage of the shock were measured on these traces at the point of maximum negative slope. Since the maximum rate of change of the light intensity occurs just behind the shock wave, the values at the maximum slope of the output will be affected by the rise time of the equipment. Assuming that the response of the equipment to a unit step function light pulse is $V=1-e^{-t'/\tau}$, it can be shown by use of the superposition integral that at any point of maximum slope, $(d^2V/dt'^2=0)$,

$$\frac{1}{I_\nu}\frac{dI_\nu}{dt'}=\frac{1}{1+\frac{\tau}{V}\frac{dV}{dt'}}\frac{1}{V}\frac{dV}{dt'}, \tag{15}$$

where I_ν is the input intensity and V is the oscilloscope deflection. τ was measured from the initial slope of the rise of the trace at the shock wave, assuming that the intensity at the shock wave was the theoretical value. The average value of τ obtained from three high Mach number runs, where equilibrium is attained rapidly, was 5.7 μsec±1.5 μsec. This is somewhat larger than would be expected from the considerations in the Experimental Arrangements section. Using the measured value of τ,

$$\left[1+\frac{\tau}{V}\frac{dV}{dt'}\right]^{-1}$$

varies from 1.2 to 4.3 in the experimental range. The measured value of the rate of energy loss is, therefore, quite sensitive to the value of τ as would be expected by the fact that τ determines the time at which the measurement is made.

In Fig. 10 the measured characteristic cooling time has been plotted as a function of temperature and compared with the theoretical curve, using Q from Eq. (12) and the density immediately behind the shock. The temperature for the experimental points was obtained by using the theoretical cooling curves (Fig. 9) and the measured value of the time at which the slope measurement was made. The largest change in temperature from the temperature behind the shock, i.e., the largest temperature correction, was 1500°K. The slowly varying term of Eq. (13) was also evaluated from the theoretical cooling curves.

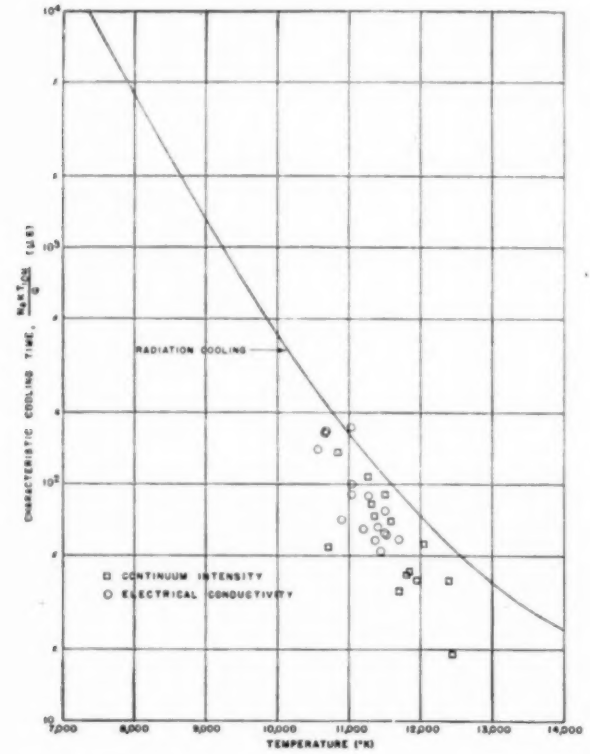


FIG. 10. Comparison between spectrophotometric and electrical conductivity measurements of the characteristic cooling time of the ionized gas and theoretical expectations assuming that the energy loss is due to continuum radiation. The agreement indicates that this is the dominant energy loss. (The density is defined by $\rho_1 = 1$ cm Hg.)

Measurements of the characteristic cooling time from the line shifts can be made by noting that in Eq. (1) the line shift is proportional to the 1.06 power of the ion density and varies only slightly with the temperature. Neglecting the dependence on temperature,

$$\frac{1}{N_e}\frac{dN_e}{dt'}=\frac{0.94}{\Delta\lambda}\frac{d\Delta\lambda}{dt'}. \tag{16}$$

The line shifts were measured from densitometer records taken at the shock and 15 μsec later. The derivative was evaluated from the difference of the line shifts and divided by their average value. The results of these measurements agreed in order of magnitude with the spectrophotometric measurements but scattered considerably. This might be expected from the fact that the characteristic cooling time depends directly on the difference of two measured line shifts, each of which have an uncertainty of the same order as the expected difference. Since considerable improvement in technique would be required to make these results of comparable accuracy with the other measurements, they have not been plotted.

The electrical conductivity of ionized argon produced by strong shock waves has been measured by Lin, Resler, and Kantrowitz[15] from the interaction of the moving gas with a magnetic field. They obtain excellent agreement between theory and experiment. The rate of change of conductivity is measured directly from the deflection of the oscilloscope traces in this experiment. We may write

$$\frac{d\sigma}{dt'}=\frac{d\sigma}{dT}\frac{dT}{d\log N_e}\frac{1}{N_e}\frac{dN_e}{dt'}. \tag{17}$$

Over the temperature range where cooling measurements were made, $d\sigma/dT$ can be taken as a constant and

$$\frac{dT}{d\log N_e}\approx\frac{T^2}{2T_{\text{ion}}}.$$

Since this varies only slightly in the experimental range, the measured characteristic cooling time depends in an essential way only on the directly measured quantity. The experimental results are shown in Fig. 9 with the temperatures evaluated as in the spectrophotometric data. The largest temperature correction was 2200°K.

The spectrophotometric and conductivity data agree within the scatter. The line-shift measurements (not presented) were somewhat higher than the others as might be expected from the crude measurement of the derivative. All of the measurements are in agreement with theory within experimental error, indicating that the dominant heat loss is due to continuum radiation. It should be noted that the somewhat uncertain temperature corrections introduced by the finite response times were ~1000°→2000°K so that the systematic departures of the spectrophotometric and conductivity data from theoretical expectations are probably not significant.

CONCLUSIONS

After taking precautions to maintain moderately high purity the visible radiation from argon, heated by a strong shock wave, consisted of argon lines which were broadened and shifted (up to 4 A) and a continuum.

Interpreting the line broadening and shift as due to Stark effect produced by random electric fields resulting from the high ion concentration in the hot gas, the line shift was employed as a technique for measuring the ion concentration utilizing a theoretical treatment of this problem by Baranger.

Measurements based on this interpretation showed that the argon reached ionization equilibrium in less than 12 μsec from the time it became luminous for temperatures greater than 11 000°K (density before the shock appropriate to 1 cm Hg pressure and room temperature).

Visible continuum radiation has been measured photoelectrically at three wavelengths for temperatures

[15] Lin, Resler, and Kantrowitz, J. Appl. Phys. **26**, 95 (1955).

from 10 000°K to 13 000°K. In agreement with theoretical expectations the intensity (per unit frequency range) was independent of frequency. Measurements of the temperature variation of the absolute magnitude of this radiation agreed with theoretical expectations and permitted us to fix a reasonable value for an undetermined factor in the theoretical intensity.

Data on the rate of cooling in the hot argon have been obtained from the variation of the continuum radiation with time, from the variation of the line shift with time, and from studies of the electrical conductivity of the gas. These data are in agreement and show that the dominant heat loss under our experimental conditions is the continuum radiation.

We acknowledge with thanks the help of Dr. E. L. Resler and Dr. S. C. Lin, who took the early Series I spectrograms. We also profited greatly by discussions of these problems with Dr. H. A. Bethe, Dr. S. Bauer, Dr. M. J. L. Baranger, and Dr. F. J. Dyson. This work was sponsored by the U. S. Office of Naval Research.

JOURNAL OF APPLIED PHYSICS VOLUME 26, NUMBER 1 JANUARY, 1955

Electrical Conductivity of Highly Ionized Argon Produced by Shock Waves*

SHAO-CHI LIN, E. L. RESLER,† AND ARTHUR KANTROWITZ
Cornell University, Ithaca, New York

(Received January 29, 1954)

Shock tube techniques for the production of shock waves up to Mach number 20 have been developed and reported previously by Resler, Lin, and Kantrowitz, J. Appl. Phys. **23**, 1390 (1952). These techniques can produce high temperature gas with accurately known enthalpy (e.g., in argon 25 percent ionization has been produced following an incident shock). Spectroscopic studies of high temperature argon produced this way by Petscheck, Rose, Glick, Kane, and Kantrowitz, "Spectroscopic studies of highly ionized argon produced by shock waves," J. Appl. Phys. **26**, 83 (1955), showed that equilibrium ionization can be reached in the time available in these experiments (of the order of 100 microseconds). This paper reports a study of the electrical conductivity of high temperature argon produced by shock waves.

At low degrees of ionization (less than 10^{-3} for argon), the diffusivity of electrons and thus the gas conductivity is determined by the cross section for electron-atom collisions which has been measured by mobility and by scattering techniques. At high degrees of ionization (larger than 10^{-3} for argon) the diffusion of electrons is primarily limited by long range Coulomb interaction with positive ions and thus is independent of the chemical nature of the gas. Theoretical treatments of this case have been given by Chapman and Cowling, Cowling, and by Spitzer and Härm. At intermediate degrees of ionization additive effects of both of these resistivity mechanisms would be expected.

Preliminary measurements with electrodes indicated large surface resistances. These effects were avoided by the development of an electrodeless technique in which the moving ionized gas deflected a magnetic field. Resultant voltages induced in a search coil were related to the conductivity distribution in the gas following the shock wave. At temperatures greater than 8000–10 000°K (depending on the gas density) the gas conductivity quickly reached a maximum value (up to 80 mhos/cm). The maximum conductivity obtained at these high temperatures agreed within 10 percent with theoretical expectations. It also agreed well with measurement of electrical resistivity in the cesium discharge by F. L. Mohler, Bur. Standards J. Research **21**, 873 (1938). At lower temperatures the oscillograms indicated that the conductivity was still rising at the end of the hot region. Under these conditions maximum conductivities reached were much lower than the theoretical values. The ionization rate obtained varied considerably with the gas density.

At the highest temperatures the conductivity declined quickly from the maximum value and the rate of decline could be correlated with the expected cooling due to recombination radiation. Indications of a high conductivity associated with luminous shock fronts were obtained.

I. INTRODUCTION

THIS paper reports the second of a series of studies of highly ionized argon produced by shock waves. This series utilizes the technique for the production of strong shock waves in shock tubes reported by Resler, Lin, and Kantrowitz.[1] The preparation of gas samples whose enthalpy is accurately known and large enough to produce a high degree of ionization is thus possible.

The first paper in the current series[2] reported a spectroscopic study of the visible radiation. This work made it clear that excitation and ionization of argon dominated the radiation only if the temperature was very high (greater than 10 000°K) or if very high purity was maintained. Measurements of the Stark effect were utilized to determine the ion density and indicated that thermal equilibrium was reached quickly (in less than 10 microseconds following the onset of luminosity) for temperatures greater than 11 000°K. A strong recombination continuum was found and absolute measurements of its intensity were made. It was shown that the continuum intensity could be correlated with a previous theoretical treatment by Unsöld.

The electrical conductivity of the gases produced by strong shock waves is comparatively easy to measure

* This work was largely supported by the Office of Naval Research.

† Now at the University of Maryland.

[1] Resler, Lin, and Kantrowitz, J. Appl. Phys. **23**, 1390 (1952).

[2] Petschek, Rose, Glick, Kane, and Kantrowitz, J. Appl. Phys. **26**, 83 (1955).

(i.e., compared with other transport properties). At low degrees of ionization conductivity measurements can be used to find the ion density in the gas. Thus, for example, the rate of approach to equilibrium ionization can be studied over quite a wide range of temperatures. At high degrees of ionization (larger than 1 percent) the diffusivity of electrons is determined by deflections due to the electric fields of the positive ions. Under these conditions it has been shown[3,4] that the electrical conductivity is independent of the nature of the gas. A conductivity measurement on a given highly ionized gas sample at known temperature and density can therefore be used for immediate comparison with other measurements using different gas samples. It also provides a direct check on the theory since no unknown parameters are involved.

The electrical conductivity of highly ionized gases plays an important role in high temperature gas dynamics (especially for large dimensions as in celestial gas dynamics) through the interaction of magnetic fields with the gas motion (magnetohydrodynamics). It will be of interest to the engineers because any information acquired concerning electron diffusivity will also be applicable to electronic heat conduction and other transport problems. The electrical conductivity of gases is of course important in gas discharge problems. In this respect, it is felt that even though much can be learned from the study of gas discharges, yet the present technique does provide another means of studying the electrical and other properties of high temperature gases under somewhat different conditions (e.g., the electron temperature in a high temperature gas produced by shock wave can be expected to be slightly lower than or approximately equal to the gas temperature, depending on the time available for adjustment; while on the other hand, the electron temperature in a gas discharge is always higher than the gas temperature). It is possible to obtain comparatively uniform and field-free samples of high temperature gases by the shock tube technique. Furthermore, because of its high time-resolving power, this technique is also suitable for the study of rate processes under gas-kinetic conditions.

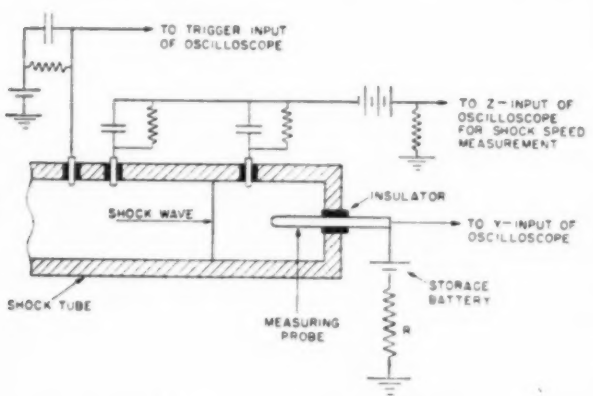


FIG. 1. Schematic diagram of electrical arrangements for the preliminary probe conductivity measurements in the shock tube. Results indicated that the probe resistance was dominated by surface effects.

II. REVIEW OF THE PROBE EXPERIMENT

Our first attempt to ascertain the degree of ionization behind a shock wave was made with a probe using apparatus schematized in Fig. 1. Measuring the voltage drop across the ionized gas with an oscilloscope is sufficient to determine the ratio of the resistance offered by the ionized gas in the annular space between the probe and the shock tube wall to the known resistance R. As the voltage across the gas was measured as a function of time the gas conductivity before and after the shock wave reflected from the end of the shock tube could be distinguished. The conductivity after reflection of the shock wave was taken to be indicated by the minimum voltage drop across the gas as measured on the oscillogram.

For small degrees of ionization, it can be shown by theory [see Eqs. (6) and (7) in Sec. VII] that the electrical conductivity of the gas is approximately given by

$$\sigma = \text{constant} \times \frac{T^{\frac{1}{2}}}{(\rho)^{\frac{1}{2}}} e^{-q_1/2kT} \tag{1}$$

where ρ is the density of the gas, T is the absolute temperature, q_1 is the first ionization potential of the gas, and k is the Boltzmann constant.

The probe experiments showed that at a given initial density, the indicated conductivity σ varied exponentially with temperature as expected for small degrees of ionization. However, the indicated conductivity was as much as a thousand times smaller than the theoretical value. Experiments run at different initial densities but at the same shock Mach number, that is, (nearly) the same gas temperature, did not give the correct density dependence. For a range of experiments where the desity was varied from 0.01 to 1 N.T.P., it was found that the conductivity was roughly proportional to the density instead of inversely proportional to the square root of the density as predicted in Eq. (1). In all cases the experimental σ was lower than the calculated value so that the equilibrium conductivity was apparently not attained. In addition the gas next to the probe and the shock tube walls is cooler than the gas away from these surfaces. It is to be expected that these cool boundary layers would greatly increase the apparent gas resistance. Rather than concentrating our attention on the complicated boundary layer effects, it was decided to design an experiment to measure conductivity undisturbed by surface effects.

[3] S. Chapman and T. G. Cowling, *The Mathematical Theory of Non-Uniform Gases* (Cambridge University Press, London, 1952), Sec. 18.12 and 14.12, and Table 6, Sec. 10.53.
[4] L. Spitzer and R. Härm, Phys. Rev. **89**, 977 (1953).

III. MEASUREMENT OF ELECTRICAL CONDUCTIVITY BY THE MAGNETIC METHOD

To eliminate the surface effect which is inherent in the probe measurements, an experiment utilizing the interaction between a magnetic field and the conducting gas behind the shock wave was designed. As shown in Fig. 2, an axially symmetric magnetic field was introduced into a glass section of the shock tube. The magnetic field was provided by a simple dc coil (or short solenoid) the axis of which was aligned to coincide with the axis of the shock tube. A small search coil was placed slightly ahead of the field coil to pick up the electromagnetic disturbance produced by the passage of the shock wave through the magnetic field.

Direct computation of the gas conductivity from the coil geometry and known gas velocity is possible in principle. It is however much easier to experimentally obtain the response of the apparatus to a moving metallic rod (or slug) which nearly similates the conducting gas geometry. This calibration will serve to evaluate a geometrical function which is difficult to calculate directly. The metal does not exactly simulate the gas in that the gas is moving relative to the shock front but this difference is easily taken into account in the analysis. An oscillogram of the search coil voltage obtained in such a calibration is given in Fig. 3.

A description of this method of correlating the gas conductivity with the metallic rod conductivity is given in Appendix I. It is seen from the results [Eqs. (A–8) and (A–10)] that the conductivity distribution $\sigma(\xi)$ can be obtained by solving an integral equation of the first kind with the metallic rod response function $V_c(s-\xi)$ as the kernel. While the direct method of solving the integral equation is not simple, an approximate method of comparison will be used for our data reduction. This method consists of first guessing the form of the conductivity distribution (simple function

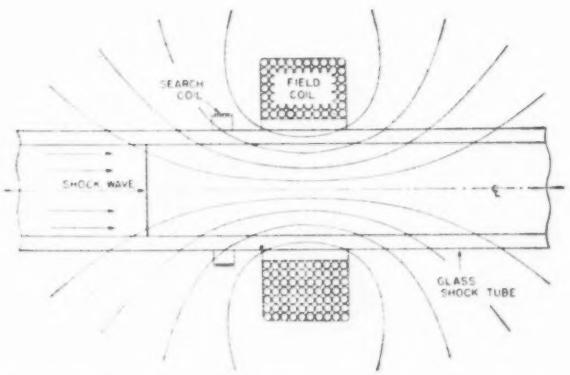


FIG. 2. Schematic diagram of the magnetic experiment designed to avoid the surface effects in the probe experiments. A magnetic field is created by the field coil. When the shock wave passes through the field coil this field is displaced, some of the lines cutting across the search coil. Since the field displacement depends upon the conductivity of the moving gas, it can be used to measure the conductivity. The apparatus is calibrated by moving an aluminum slug of known conductivity and known velocity through the shock tube. A detail of the search coil is given in Fig. 4.

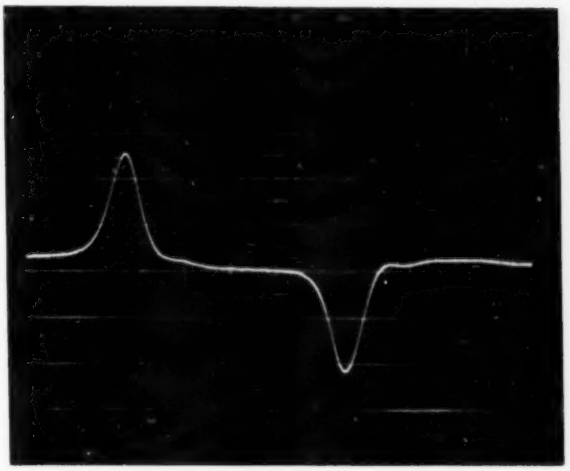


FIG. 3. Calibration oscillogram made with the apparatus of Fig. 2 by passing an aluminum slug through the shock tube at 780 cm/sec. This oscillogram gives essentially the response of the apparatus to positive and negative conductivity step functions. Calculated response for some other types of conductivity variations are shown in Table I.

with adjustable parameters) and then calculating the voltage curve according to Eq. (A–8). Comparison between this calculated voltage curve and the actual observed voltage signal enables the determination of the important parameters of the conductivity distribution, such as the maximum conductivity, the approximate rate of rise of conductivity during the build-up period, the rate of decrease of conductivity due to cooling, etc. The inaccuracy introduced by the assumed conductivity function is probably not the largest source of error.

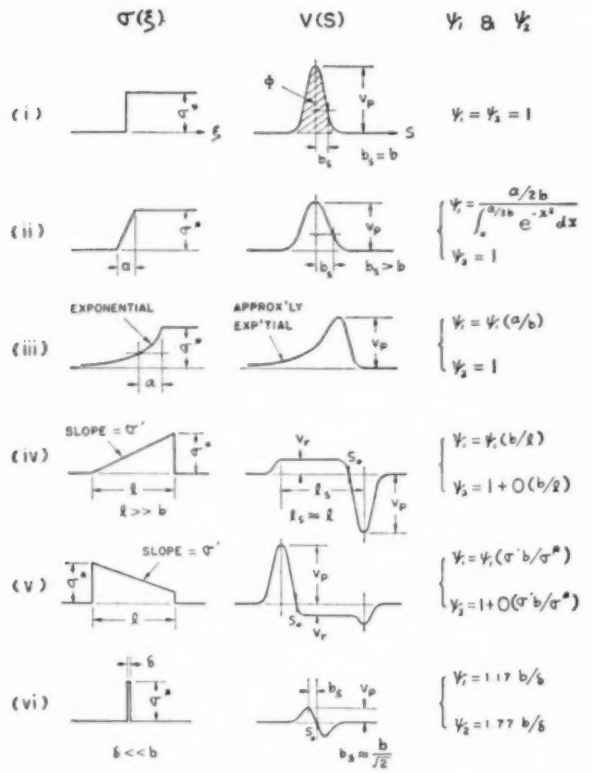
To facilitate the interpretation of the experimental results presented later in Sec. VI, a summary of the calculated voltage curves due to various simple conductivity distributions is schematically presented in Table I. In all cases, it is possible to express the ratio between the maximum conductivity σ^* attained by the gas behind the shock wave and the metallic rod conductivity σ_c as,

$$\frac{\sigma^*}{\sigma_c} = \frac{u_c^2 I_c V_p}{U u_2 I V_{cp}} \psi_1 = \frac{u_c^2 I_c \Phi}{U u_2 I \Phi_c} \psi_2, \tag{2}$$

where V_{cp} and Φ_c are the peak and the pulse area of the metallic rod response curve, respectively; V_p and Φ are the peak and the pulse area of the calculated voltage curve as shown in Table I [for cases (iv) and (v), the pulse area is taken up to the zero-voltage point s_0]; ψ_1 and ψ_2 are functions dependent on the conductivity distribution and can be evaluated numerically from Eqs. (A–8) and (A–9); and other symbols are the same as were defined in the Appendix.

From Fig. 3, it is seen that the response pulse for the metallic rod can be closely approximated by a Gaussian

TABLE I. Search coil response $V(s)$ for various conductivity distributions $\sigma(\xi)$. The functions ψ_1 and ψ_2 are defined by Eq. (2). Note: $0(\epsilon)$ denotes a term of the order of ϵ.



function so that

$$V_c(s) \approx V_{cp} \exp -[(s-s_p)/b]^2, \tag{3}$$

where s_p is the position of the pulse peak and b is a characteristic length which is a measure of the effective width of the magnetic field and is also a measure of the resolving power of this experiment for the conductivity distribution. Using this approximation, the functions ψ_1 and ψ_2 can sometimes be expressed in closed form such as in case (ii) of Table I. At any rate, it should be pointed out that ψ_2 will be unity for the cases (see Table I) in which there is a sufficiently long region of constant conductivity, and that it will be unity plus a small correction term for the cases where the variation of conductivity in a distance b is small. Therefore, the pulse area relation is more conveneint for the determination of maximum conductivity behind the shock wave, while the pulse peak relation can be used to find the rate of conductivity buildup after σ^* has been determined.

From Eq. (A-8), it is seen that the slope of the conductivity distribution in (iv) and (v) of Table I is given by

$$\sigma' = \frac{u_c^2 I_c V_r}{U u_2 I \Phi_c} \sigma_c, \tag{4}$$

where V_r is the voltage of the flat region of the $V(s)$ curve.

IV. APPARATUS

(1) Description of Shock Tube

The techniques of Resler, Lin, and Kantrowitz[1] were used for the production of strong shock waves. The shock tube and gas purity techniques employed for these experiments were the same as those used in taking the third series spectra[2] (except for the preliminary probe measurements which employed the shock tube and purity techniques of the second series spectra of Petschek, Rose, Glick, Kane, and Kantrowitz).[2] In the third series spectra taken at shock Mach numbers higher than 11 (about 10 000°K), only argon lines appeared confirming the indicated high purity (better than 99 percent). For all the experiments with the exception of the last half of the $p_1 = 10$-cm Hg series, the driver section was 19 in. long. For the last half of the 10-cm series, the driver section was extended to 27 in. in an effort to improve the uniformity of the shock strength. The low pressure section of the shock tube consisted of a 4-ft steel pipe at the upstream end (adjacent to the high pressure driver section), a 5-ft glass pipe of $1\frac{1}{2}$-in. inside diameter in the middle, and then a 2-ft steel pipe at the downstream end connecting the glass pipe to the dump chamber.[2] The end of the low pressure section was separated from the dump chamber by a thin celluloid diaphragm, and the inside of the two steel pipes has been nickel-plated to improve the gas purity.

Conductivity measurements using the magnetic method were made for three different initial argon densities, namely: 0.1-, 1-, and 10-cm Hg pressures at room temperature. Of these three initial densities, the 1-cm pressure was found most suitable for our apparatus and therefore more experimental data was obtained for this series than the other two.

The 0.1-cm series suffered the disadvantage of high relative impurity in the argon since the shock tube could not be evacuated to pressures much lower than 0.01-cm Hg with the existing vacuum system. Therefore, only the high Mach number points for this series were of any significance. On the other hand, for the 10-cm series, the Mach number range was limited because the glass section could not withstand strong shock waves with this initial density (the lower limit of usable shock strength for these experiments was determined by the sensitivity of the detector, that is, the search coil and the vertical amplifier of the oscilloscope).

(2) Shock Speed Measurement

The shock speed measurements for these experiments were accomplished by taking rotating-drum camera pictures of the shock waves propagating down the shock tube, through a slit (about 1 mm wide) parallel to the shock tube axis in the glass section. Knowing the drum speed and the image-to-object ratio of the camera, the shock speed can be computed from the slope of the luminous shock front.

It should be pointed out that this method of shock speed measurement is ideal only when the shock front itself is luminous[1] and can be clearly identified in the drum camera picture. It was found however, that the luminous shock front was observed only occasionally (the actual cause of appearance and disappearance of the luminous shock front is not yet known). Therefore, whenever the position of the shock front was not well defined, the shock speed was obtained indirectly by measuring the luminous particle paths trailing behind the shock wave. This method was considered satisfactory since the relation between the particle speed and shock speed has been found quite consistent from those drum camera pictures where both speeds can be measured at the same time. (It was found that the observed particle speed was always about 7 percent faster than the theoretical particle speed u_2 given in the Appendix of reference 1, but the discrepancy was quite consistent.)

Another complication in shock speed measurement arose when the shock strength was not uniform as the shock wave travelled down the shock tube, because the thermodynamic state of the hot gas depends to a certain extent on the past history of the shock wave. This has been found to happen for some experiments where the $O_2-H_2-N_2$ combustible mixture was used as the driver gas. In some cases, it was found that delayed combustion in the driver mixture after the diaphragm burst set up compression waves in the conducting region and hence tended to accelerate the shock wave. In another case (especially for the high initial argon pressure experiments where the wave speed in the driver mixture is sufficiently high compared with the shock speed), the expansion waves from the far end of the driver section caught up with the hot region and hence slowed down the shock wave. Therefore, even though the overall inaccuracy of the shock speed measurement (due to inaccuracy in drum speed measurement, slit alignment, slope determination, etc.) was kept to within about ± 2 percent when the shock front was well defined, the actual inaccuracy when nonuniformity in shock strength is present may be considerably larger.

(3) Triggering and Synchronization

To obtain the search coil voltage history when the shock wave passes through the magnetic field, a single sweep from a Tektronix 512 oscilloscope was used for each experiment. The sweep was triggered by an RCA 931-A photomultiplier tube located somewhere upstream of the search coil to detect the approach of the shock wave. The gate output of the oscilloscope was used in turn to trigger a small synchronizing spark which registered on the drum camera picture. From the location of the spark image on the drum camera, the horizontal scale of the oscillogram can be correlated with the position of the luminous region relative to the search coil.

(4) Field and Search Coils

The magnetic field coil mentioned in the Appendix consisted of 800 turns of No. 18 (0.040-in. diam) enamelled wire wound on a Bakelite coil-form of such dimensions that the resultant solenoid has a length of 3.8 cm and a mean diameter of 8.6 cm. The Bakelite coil-form had a bore of 5.4 cm so that it could be fitted into the glass section of the shock tube allowing about 3-mm clearance between it and the outside surface of the glass pipe. The whole coil assembly was then supported on a platform which was completely isolated from the shock tube. The search coil was also mounted on the coil-form that carried the field coil, and therefore it was also mechanically isolated from the shock tube. The range of the dc field current was between 6 and 16 amperes giving axial field strengths of 3000 to 8000 gauss.

The search coil was a 50-turn 8 mm long single-layered coil made of No. 36 (0.005-in. diam) enamelled wire. As shown schematically in Fig. 4, the search coil was actually made up of two 25-turn coils wound side by side and connected in series. The junction point of the two coils was grounded, and the two ends of the combined coil were connected to the push-pull input of the Tektronix 512 oscilloscope. The purpose of this center-tap arrangement was to minimize the electrostatic pickup which will be discussed in Sec. V. The value of the shunt input resistance (1000 ohms) was so

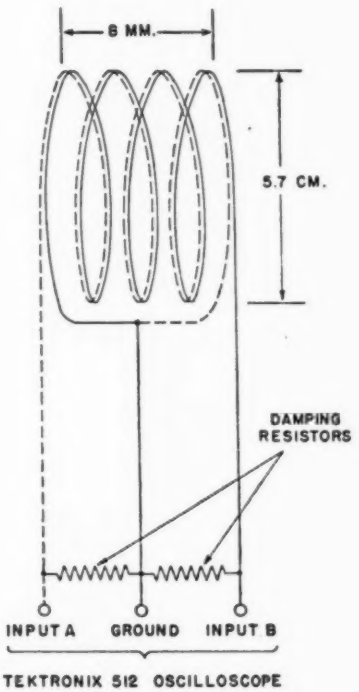


FIG. 4. Schematic diagram of the search coil used in the apparatus of Fig. 2. The centertapped design is intended to minimize electrostatic pickup due to the sudden appearance of (slightly) charged gas in the shock tube (see Sec. V).

chosen to give approximately critical damping of the search coil circuit.

As mentioned before, the search coil was mounted on the front section of the Bakelite coil-form with a 14-mm gap between it and the field coil. The whole magnetic field coil and search coil assembly was located approximetely one foot ahead of the downstream end of the glass section.

(5) Calibration

The metallic rod employed to obtain the calibration voltage curve $V_c(s)$ mentioned in Sec. III was an aluminum alloy rod 18.9 cm long with a diameter approximately 0.3 mm smaller than the inside diameter of the glass section. This rod was machined from a long 24 *ST* Dural bar stock the electrical conductivity σ_c of which has been measured to be 1.82×10^5 mhos/cm at room temperature. The rod was catapulted through the magnetic field (with the glass pipe as a guide) by stretched rubber bands attached to a small lug at the front end of the rod. The length and position of the catapult were so adjusted that the rubber bands exert no force when the rod enters the magnetic field. The approximately uniform speed of the rod thus obtained was measured from the time elapsed between the occurrence of the front and rear pulse in the oscillogram (see Fig. 3). The average value of the quantity $u_c^2 I_c/V_{cp}$ from ten such experiments was approximately 6.9×10^7 (cm/sec)^2-amperes per volt, with a probable inaccuracy of about 10 percent. The average value of the effective width b of the magnetic field as defined in Eq. (3) was found to be approximately 1.8 cm.

The value of the parameter $\sigma_c u_c$, which is a measure of the skin depth of the moving conductor (see Appendix I), was varied between 2.3×10^7 and 1.4×10^8 mhos/sec, corresponding to 0.5 and 3 times the σU for the strongest shock wave. No appreciable variation of the quantity $u_c^2 I_c/V_{cp}$ (less than 10 percent) and of the form of the response curve was observed throughout this range of $\sigma_c u_c$, confirming that the skin effect was small for our experiments.

V. ELECTROSTATIC EFFECT

Before going into the discussion of the conductivity results, we should mention something about the electrostatic phenomenon associated with the shock wave in a glass shock tube. The electrostatic effect was sometimes so prominent that unless proper precautions were employed to suppress or to identify it, one would quite easily be misled into erroneous interpretation of the experimental results.

In the early probe experiments, it was noticed that when an electrically floating probe was directly connected to an oscilloscope, a large voltage signal was picked up when the shock wave passed by the probe. The voltage seemed to be higher for the weaker shock waves where the conductivity of the gas was lower.

During the early stages of the magnetic experiment, where a simple 50-turn search coil (connected to the single-sided input of the oscilloscope) was used, electrostatic effects were also noticed. For example, during one experiment where the magnetic field was purposely left off to test the noise level of the search coil circuit, an unexpectedly large signal was observed when the shock wave passed through the search coil. In addition, when the shock wave reached the end of the glass pipe and came into contact with the steel section, another signal of opposite polarity to the first one was picked up. The duration of these signals was of the order of 5 microseconds or about half an inch of shock wave displacement. Later, an experiment was performed in which both ends of the simple search coil were soldered together and connected to the oscilloscope input (the metal portions of the shock tube being grounded to the scope), larger signals were observed (of the order of 0.1 volt and lasting about 30 microseconds). The last experiment showed that these signals were electrostatic (and not magnetic) pickups due to the finite capacitance between the search coil and the gas inside the glass shock tube. Thus, if there was a net charge in the electrically isolated conducting gas behind the shock wave, there will be a signal when the gas first moves into the vicinity of the search coil, and then there will be another signal when the charge in the gas is grounded through the steel section of the shock tube.

Further examination of the signal of the last experiment indicated that the form of the signal was not quite the same as that expected from a uniformly charged body passing through the search coil. Therefore it suggested that the charge distribution in the gas was not uniform. This incidentally is a case that would be expected in a gas where the ion concentration is not uniform so that the electrons tend to diffuse toward the region of low concentration thus setting up an electric field.

While the study of the electrostatic behavior of the gas appeared to be a promising technique which could be developed to yield useful information concerning the properties of the ionized gas behind shock waves, it has not been carried any further because our attention was directed to the measurement of electrical conductivity. To this end, the center-tapped search coil described in the last section was developed. It is seen in Fig. 4 that the capacitive pickup from the two ends of this coil will be canceled out by the push-pull arrangement, while the magnetic pickup was unaffected. Even though it is true that the electrostatic effect cannot be completely eliminated by this method due to imperfect balance of the circuit (for instance, the finite dimensions of the coil would introduce a phase difference among the different parts of the search coil, which would be difficult to compensate), the residual signal has been found so small and of so short duration that it can be easily distinguished from the main magnetic signal of the experiment. Therefore, the arrangement of this search coil circuit was considered satisfactory.

VI. EXPERIMENTAL RESULTS

In Fig. 5 are shown four typical oscillograms together with the accompanying drum camera pictures for the 1-cm series at four different Mach numbers. Let us examine these pictures in sequence of increasing shock strength.

(1) Low Mach Number

Figure 5(a) shows the results of a low Mach number experiment ($M=7.5$) in which the shock front was luminous. It is seen in the drum picture that the gas behind the shock front was practically nonluminous except for the rear region which contained gas particles initially in the upstream steel section before the diaphragm breakage. The voltage signal similarly indicated a finite conductivity at the shock front (probably more than three times the maximum conductivity attained by the gas behind the shock wave; more discussion on luminous front conductivity will appear later in this section), which was followed by a region of very small conductivity (less than 0.01 mho/cm), and then in the region further back the conductivity was measurable again. In the oscillogram, there was also a small high frequency (short duration) disturbance in the rear signal which is believed to be the residual electrostatic pickup mentioned in the previous section. This disturbance was visible only in the very low Mach number oscillograms where maximum vertical amplification was used.

(2) Intermediate Mach Number

Figure 5(b) shows the drum camera picture and the oscillogram for a shock wave somewhat stronger ($M=10.6$) than the previous one. From the drum camera picture, it is seen that the gas luminosity behind the shock wave built up gradually and that the luminosity of the shock front was negligible compared with the gas luminosity far behind. According to the analysis (see Table I) it is seen from the oscillogram that except for small fluctuations the gas conductivity behind the shock wave was building up almost at a constant rate until it was cut off by the contact discontinuity. Careful examination of the small fluctuations of the voltage signal indicated that they were oscillations of more or less constant frequency (about 8×10^4 cycles per second) Approximate calculation showed that this frequency was roughly the same as that of the first mode of radial oscillation of the high temperature gas inside the $1\frac{1}{2}$-in. i.d. glass pipe. Therefore this fluctuation of the voltage signal was probably due to the interaction between radial oscillation of the conducting gas and the axial component of the magnetic field. A rough analysis based on this interpretation gave radial velocities of the order of 2 percent of the axial gas velocity. Since the period of this voltage fluctuation was sufficiently small compared to the duration of the entire voltage signal, the mean line of the oscillating voltage may be taken as the signal essentially due to the axial motion of the gas, and thus the results of the previous analysis were still applicable.

(3) High Mach Number

As shown in Fig. 5(c), at sufficiently high Mach number (about 12), the electrical conductivity behind the shock wave built up rapidly to a constant value and stayed at this value unitl it was cut off by the contact discontinuity. Therefore, the resulting oscillogram was very similar to that produced by a metallic rod (Fig. 3). In the drum camera picture, the liminosity of the gas was correspondingly very uniform.

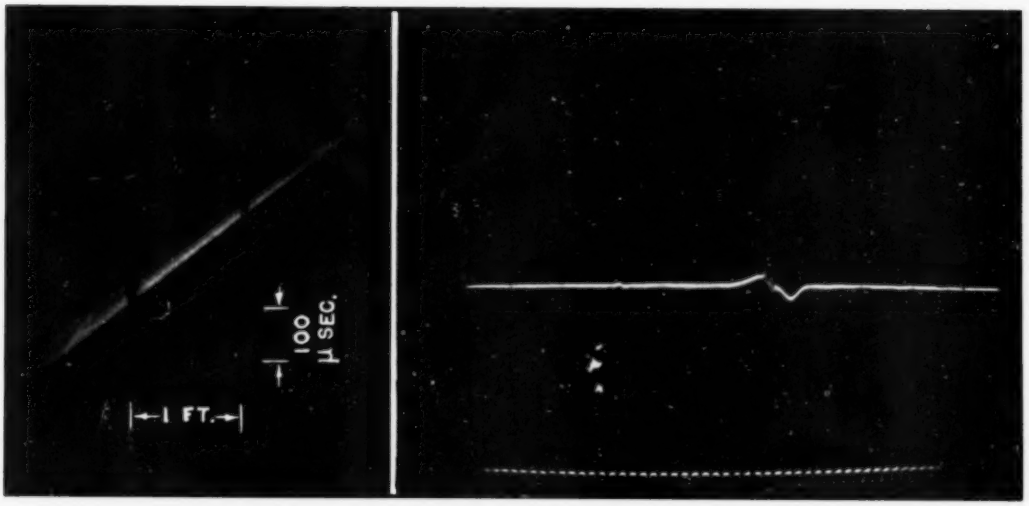
(4) Very High Mach Number

At still higher Mach numbers, the temperature behind the shock wave will be so high that the radiative heat loss becomes a controlling factor for the thermodynamic state of the highly ionized gas (see reference 2, Fig. 5). As shown in Fig. 5(d), where $M=16.4$, the conductivity of the gas rose very rapidly to a maximum value and then began to decline at a rate which was decreasing toward the contact discontinuity according to the analysis of Sec. III [see Table I(v)]. The drum camera picture also showed a corresponding trend in gas luminosity.

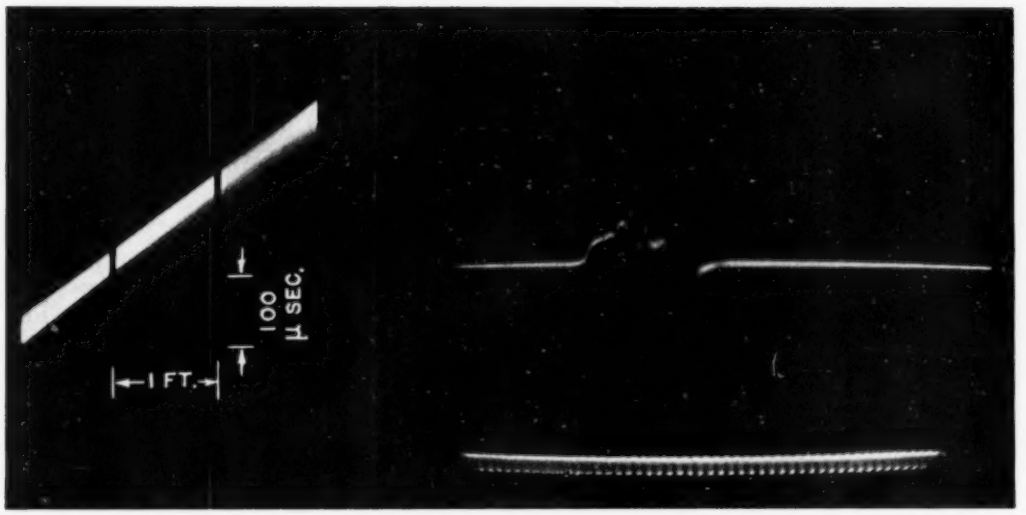
Maximum Conductivity Results

The maximum electrical conductivity attained by the argon gas behind the shock wave for the three series of experiments (0.1-, 1-, and 10-cm Hg initial pressures) was plotted in Figs. 6(a), (b), and (c) for comparison with conductivity theories. These experimental points were computed from the time integrals of the voltage with appropriate first order corrections [see Table I, (iv) and (v)] for finite buildup distance, cooling effect, etc. These corrections were generally less than 20 percent of the uncorrected value.

The temperature scale for the experimental points was obtained from the temperature-Mach number relationship given by Resler, Lin, and Kantrowitz.[1] Since the shock speed and not the gas temperature was actually measured during each experiment, the temperature scale is, strictly speaking, an enthalpy scale. This scale will give correct temperature reading only when the gas behind the shock wave has reached the equilibrium degree of ionization. It was found, however, both from the spectroscopic study by Petschek, Rose, Glick, Kane, and Kantrowitz,[2] and from the present experiment, that at high temperatures, thermal equilibrium was reached quickly behind the shock wave. On the other hand, at low temperatures, the equilibrium degree of ionization was so low that the amount of energy required for ionization was small compared with the kinetic energy of the gas. Therefore, at low temperatures the enthalpy scale is again a good representation



(a)



(b)

FIG. 5(a), (b), (c), (d). Typical oscillograms for the magnetic experiment with the accompanying drum camera pictures (on the left side). The timing markers are 10 micorseconds apart. The shock Mach number M, the equilibrium temperature T_2, the oscilloscope vertical sensitivity (the horizontal lines on the screen were 1 cm apart), and the field coil current for these four experiments are as follows:

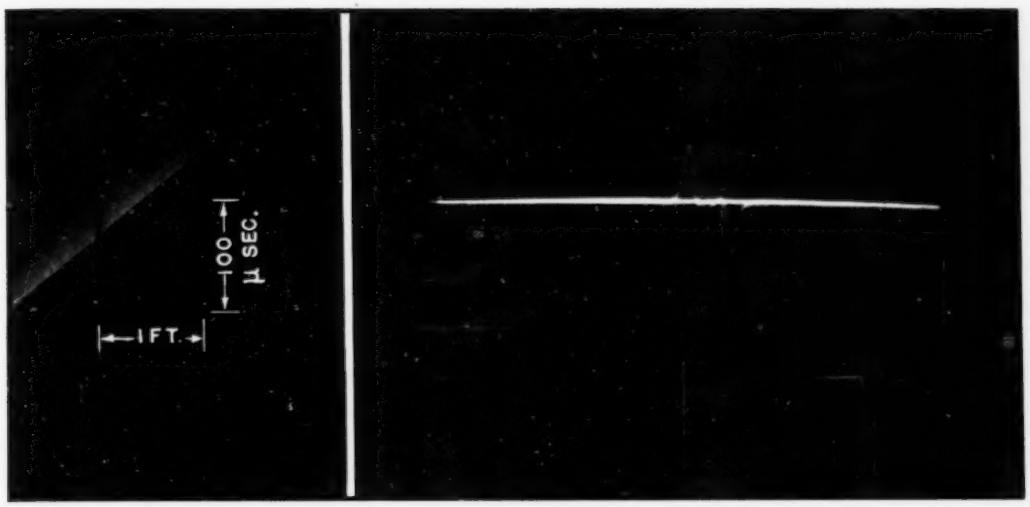
	(a)	(b)	(c)	(d)
M	7.5	10.6	12.5	16.4
T_2, °K.	5500	9800	11 300	13 500
Sensitivity, volts/cm	0.010	0.40	5.0	3.0
Field current, amperes	10.4	9.7	15.2	6.7

On the drum camera pictures, the shock wave traveled from left to right while the time axis was directed upward. The synchronizing spark, which is not visible on these plates, was recorded on the extreme left of the drum pictures. Maximum conductivity points obtained from these four oscillograms are identified in Fig. 6(a).

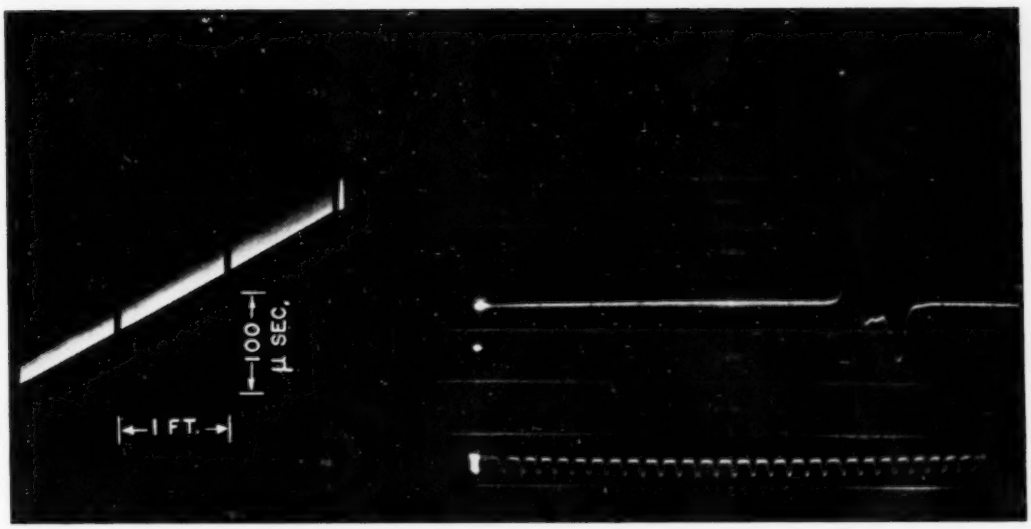
of temperature no matter whether the equilibrium degree of ionization has been reached or not.

Cooling Rate

The rate of decline of electrical conductivity behind the shock wave for the high Mach number experiments can be computed from the nearly horizontal portion of the voltage signal (see Table I(v)] according to Eq. (4). From this rate of decrease of conductivity, the rate of energy loss can be estimated from the relations among conductivity, temperature, degree of ionization and enthapy [Eq. (12), and the reference 1, Appendix], and by considering the dynamics of the gas flow. This has



(c)



(d)

been carried out and the results were expressed in terms of a characteristic cooling time and plotted in Fig. 6 of reference 2. It was found that these results agreed quite well with both the spectrophotometric measurements and the Unsöld-Kramers theory of continuum radiation (with effective ionic charge $z=1$ and energy range $\Delta E=kT$).

Luminous Front Conductivity

It has been shown in the analysis of Sec. III that if the thickness δ of a thin conducting region [such as the luminous shock front[1,5] which can be seen on Fig. 5(a)], is known, the mean electrical conductivity of the region can be calculated according to Eq. (2) with $\psi_1=1.17b/\delta$ or $\psi_2=1.77b/\delta$. In other words, the product $\sigma\delta$ is directly measurable from the voltage signal caused by the passage of the thin conducting region through the magnetic field. Thus, if we take the small first wriggle of the oscilloscope trace of Fig. 5(a) as the magnetic signal of the luminous front, and using the thickness measured by Petschek[5] (of the order of 1 mm), we obtain a conductivity of approximately 0.3 mho/cm (or about 3 times the final conductivity) for the luminous front. Since the actual thickness of the luminous front might have been somewhat thinner than 1 mm [it should be pointed out that the apparently rather thick luminous front on the drum picture of Fig. 5(a) was due to a wide slit used for this low Mach number experiment and hence it does not give an estimate of the actual thickness of the front], the actual conductivity at the luminous front might have been considerably higher than the above estimate. On the other hand, on account of the fact that the duration of the small wriggle was about a factor of two shorter than the duration expected from the analysis,

[5] H. E. Petschek, Phys. Rev. 84, 614 (1951).

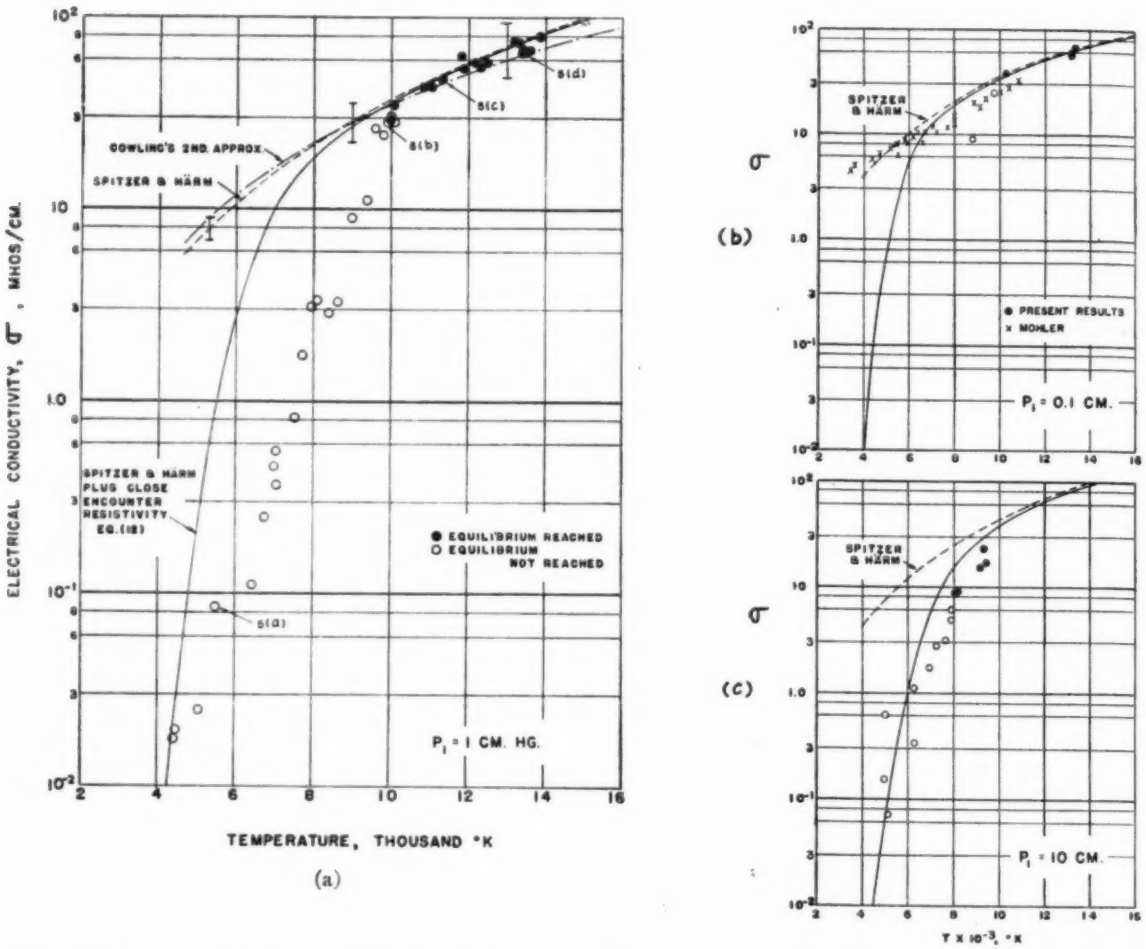


FIG. 6. Maximum electrical conductivities in argon heated by shock waves measured by the magnetic experiment of Fig. 2. The oscillograms taken at the lower temperatures [Fig. 5(a) and (b)] indicate that the conductivity was still rising at the end of the time available and the indicated maximum conductivities reached are plotted as open circles. It will be seen that the solid points [Fig. 5(c) and (d)] in which equilibrium conductivity was apparently reached are in good agreement with the theory of the highly ionized gas in spite of the uncertainty in this theory (vertical lines) pointed out by Spitzer and Härm. In the 10-cm initial pressure experiments the conductivity was lower than the theoretical value even when equilibrium was apparently reached. For these points close encounter resistivity was important and the results indicate that the cross-section taken from mobility measurements (Townsend and Bailey) and employed in the theory were too small. Comparison of the results at various initial pressures indicates that equilibrium was more readily attained at higher density. Mohler's resistivity measurements in cesium discharge at pressures from 0.0045 to 0.080-mm Hg, (taken from Fig. 1 of Reference 17) are plotted in (b) as crosses for comparison with the present results.

and from the fact that it was of the same order of magnitude as the electrostatic pickup, it was also possible that this signal was simply residual electrostatic pickup of the search coil due to the shock front (see Sec. V). It is our intention to carry out further experimental investigations of the luminous front phenomenon.

VII. COMPARISON WITH THEORY OF ELECTRICAL CONDUCTIVITY IN GASES

Theories of electrical conductivity in gases have generally been developed for either one of the two extreme cases. Namely, (1) the case of a very slightly ionized gas where the close encounters between the electrons and the neutral atoms dictates the electron mobility; and (2) the case of a completely ionized gas where the distant encounters between the ions dominate the situation. For the intermediate case where the degree of ionization is moderate (such as in the intermediate Mach number range of our experiments), no detailed treatment is known to the authors. To cover this intermediate range, the following approximation will be used,

$$\frac{1}{\sigma}=\frac{1}{\sigma_0}+\frac{1}{\sigma_d}. \tag{5}$$

In the above expression, $1/\sigma_0$ denotes the electrical resistivity entirely due to close encounters between electrons and the gas molecules and positive ions, $1/\sigma_d$ denotes the resistivity due to distant encounters between electrons and the positive ions, and σ denotes the resultant conductivity of the gas. For pure argon in the density range of our experiments, the two terms σ_0 and

σ_d will be equal at about 7000°K temperature, or a degree of ionization of about 10^{-4} (i.e., when 0.01 percent of the original gas molecules was ionized). At lower temperatures the resultant conductivity will be determined essentially by the close encounters, while at higher temperatures the conductivity will be determined by the distant encounters.

Following is a brief review of the existing conductivity theories which will be used to calculate the terms σ_0 and σ_d in Eq. (5).

(1) The Slightly Ionized Gas

The electrical conductivity of a slightly ionized gas of rigid elastic spherical molecules is given by Chapman and Cowling,[3]

$$\sigma=0.532\frac{\alpha\epsilon^2}{(m_e kT)^{\frac{1}{2}}}\frac{1}{Q}, \tag{6}$$

where $\alpha\equiv n_e/n_0$ is the degree of ionization (n_e being the number density of free electrons and n_0 being the number density of gas molecules), ϵ is the electronic charge, m_e is the mass of an electron, k is the Boltzmann constant, T is the absolute temperature, and Q is the electron-atom collision cross section. Chapman and Cowling have also worked out the diffusion coefficient (and hence the conductivity) for various molecular models with different interaction forces. But these results are not directly applicable to the electron-argon interaction because of the Ramsauer-Townsend effect.[6-8] It can be shown (Appendix II), however, that Eq. (6) is still valid provided that Q is an appropriate average of the cross section over the electron speed distribution. Since the collision cross section for argon as a function of electron speed has been measured by Ramsauer and others, Q can be computed according to Eq. (A–18). The effective cross section thus computed, using Normand's measurement[9] and Maxwellian electron speed distribution [Eq. (A–19)], is plotted in Fig. 7 (dotted curve) as a function of temperature.

The effective cross section Q can also be obtained from mobility measurements such as those of Townsend and Bailey,[7] Nielsen,[10] Wahlin,[11] and others. Unfortunately the measurements of Townsend and Bailey (1922) are indirect in our energy range and the accuracy of this work is not clear. The more recent measurements of Nielsen and Wahlin are, respectively, at energies above and below our range. The effective cross section from Townsend and Bailey's mean free path in argon is plotted in Fig. 7 (solid curve) where the temperature was taken as appropriate to the velocity of agitation of the electrons (i.e., $u^2=3kT/m_e$ where u is the velocity of agitation as defined by Townsend and Bailey.[7]) It should be pointed out that the electron swarm in Townsend and Bailey's experiment would have the Druyvesteyn velocity distribution rather than the Maxwell distribution which would be expected for all but the lowest temperature experiments reported here and therefore there should be a correction to the cross section *versus* temperature relation. The indicated accuracy of the measurements however does not justify this correction and hence no such correction was made in constructing Fig. 7.

It is seen in Fig. 7 that the effective cross section obtained from Normand's measurement is considerably higher than that obtained from Townsend's and Bailey's measurement, especially at low temperatures. Since the deflection method used by Normand (Ramsauer's method) tends to over emphasize the effect of the very small deflections, it gives too high a value for the effective cross section for diffusion. Therefore, Townsend and Bailey's mobility measurement was chosen for comparisons with our experiments.

For a slightly ionized monatomic gas, the degree of ionization α at equilibrium is (from Saha's equation),

$$\alpha=G\frac{T^{5/4}}{(p)^{\frac{1}{2}}}e^{-q_1/2kT}, \tag{7}$$

where G is a constant (depending on the statistical weights of the positive ion and the atom) and q_1 is the first ionization potential of the gas atom. For argon the value of G is 1.73×10^{-2} when T is given in °K and p in cm Hg; and the established value for q_1 is 15.69 ev.

It is seen from Eqs. (6) and (7) that when the degree of ionization is low, the electrical conductivity varies exponentially with temperature while its dependence on pressure (or density) is weak. It is also obvious that the conductivity is very sensitive to those impurities which have lower ionization potential than the gas.

(2) The Completely Ionized Gas

The Chapman-Enskog method of solving the Maxwell-Boltzmann equation has been applied by Chapman and Cowling,[12] Cowling,[13] and Landshoff[14] to obtain the first, second, and higher approximations for the electrical conductivity of completely ionized gases. More recently Spitzer and his collaborators[4,15] have solved the same problem from a different approach. They considered the effect of electron-ion encounters as a problem of diffusion in velocity space and obtained a result which was very close to the higher approximations of the Chapman-Cowling method, except for a different cut-off distance in the integration over the impact parameter.

[6] C. Ramsauer, Ann. Physik **64**, 513 (1921); **66**, 546 (1921).
[7] J. S. Townsend and V. A. Bailey, Phil. Mag. **43**, 593 (1922); **44**, 1033 (1922).
[8] H. S. W. Massey and E. H. S. Burhop, *Electronic and Ionic Impact Phenomena* (Oxford University Press, London, 1952), Chap. I.
[9] C. E. Normand, Phys. Rev. **35**, 1217 (1930).
[10] R. A. Nielsen, Phys. Rev. **50**, 950 (1936).
[11] H. B. Wahlin, Phys. Rev. **37**, 260 (1931).
[12] See reference 3, Sec. 10.33.
[13] T. G. Cowling, Proc. Roy. Soc. (London) **A183**, 453 (1954).
[14] R. Landshoff, Phys. Rev. **82**, 442 (1951).
[15] Cohen, Spitzer, and Routly, Phys. Rev. **80**, 230 (1950).

According to Spitzer and Härm,[4] the electrical conductivity of a completely singly ionized gas (i.e., each gas atom has been ionized once) is given by

$$\sigma=\frac{0.591(kT)^{\frac{3}{2}}}{m_e^{\frac{1}{2}}\epsilon^2\ln(h/b_0)}, \tag{8}$$

where h is the Debye shielding distance in the plasma and is given by

$$h^2=\frac{kT}{8\pi n_e\epsilon^2},$$

while $b_0\equiv\epsilon^2/3kT$ is the impact parameter at which a positive ion deflects a mean-energy electron by 90 degrees. As Spitzer and Härm pointed out, the above theory has a relative inaccuracy of the order of $1/\ln(h/b_0)$ as a result of the neglect of some interaction terms in the analysis. This inaccuracy amounts to about ± 15 to 30 percent under our experimental conditions.

The higher approximations of the Chapman-Cowling method gave results which differ from Eq. (8) essentially in the logarithmic term. In these cases, the term $\ln(h/b_0)$ was replaced by $\ln(d/b_0)$, where $d\equiv 4/3n_e^{-\frac{1}{3}}$ is a measure of the mean distance between neighboring particles in the gas. Under our experimental conditions, the two logarithmic terms differ from each other only by 10 or 20 percent, which is smaller than the relative inaccuracies of these theories.

It is seen from the above theories that for a completely ionized gas, the electrical conductivity is approximately proportional to the $\frac{3}{2}$ power of temperature, and it changes only very slowly with ion concentration density (through the logarithmic term). Moreover, it is independent of the chemical nature of the gas.

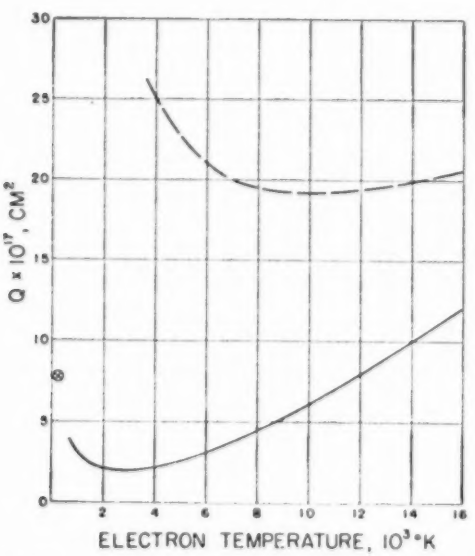


FIG. 7. Effective cross section Q of argon as a function of electron temperature. Solid curve was obtained from Townsend and Bailey's (see reference 7) mean free path results. Dotted curve was computed from Eq. (A-19) using Normand's (see reference 9) cross-section measurements (Ramsauer method). Single point at low electron energy was given by Wahlin (see reference 11).

(3) Resultant Conductivity

Considering the diffusion of electrons through the positive ions as a free-path process, then comparison between Eq. (8) and Eq. (6) shows that a positive ion has the same effect in obstructing the motion of the electrons as a particle with effective cross section

$$Q_i=8.10b_0^2\ln(h/b_0). \tag{9}$$

Therefore, in an ionized gas where there are n_a neutral atoms and n_i positive ions per unit volume with effective cross sections Q_a and Q_i, respectively, the mean-free path of the electron is given by (note that collisions between electrons have no effect on the electron mobility),[16]

$$\lambda=\frac{1}{n_aQ_a+n_iQ_i}. \tag{10}$$

Consequently, the electrical conductivity of the gas is given by

$$\sigma=0.532\frac{n_e\epsilon^2}{(m_ekT)^{\frac{1}{2}}}\frac{1}{n_aQ_a+n_iQ_i}. \tag{11}$$

If we let n_a equal to n_0, the original number density before ionization took place, to take into account approximately the effect of close encounters between the electrons and the psitive ions which has been left out in Spitzer's analysis; and assume that there is no net charge in the gas so that $n_i=n_e$, we obtain

$$\sigma=\left\{1.88\frac{(m_ekT)^{\frac{1}{2}}Q_a}{\alpha\epsilon^2}+1.69\frac{m_e^{\frac{1}{2}}\epsilon^2}{(kT)^{\frac{1}{2}}}\ln(h/b_0)\right\}^{-1}, \tag{12}$$

which is the same as taking the conductivities given in Eq. (6) and Eq. (8) as σ_0 and σ_d in Eq. (5).

The theoretical conductivity for argon behind the incident shock wave under equilibrium conditions was calculated according to Eq. (12) and plotted in Fig. 6(a), (b), and (c) for the three different initial densities at which the experiments were performed. The temperature and other thermodynamical functions such as density, degree of ionization, etc., were calculated according to the method described in the Appendix of reference 1, using the shock Mach number as a parameter. Townsend and Bailey's cross section for argon (the lower curve in Fig. 7) was used to calculate the close encounter term in Eq. (12).

By comparing the relative magnitude of the close and distant encounter terms in Eq. (12), it was found that in the density range of our experiments the cross-over point (at which the two terms σ_0 and σ_d are equal) occurs at a degree of ionization of about 10^{-4} or a temperature of about 7000°K in argon. At temperatures

[16] See for instance, J. Jean, *An Introduction to the Kinetic Theory of Gases* (Cambridge University Press, London, 1952), Sec. 171.

higher than about 8000°K ($\alpha > 10^{-3}$), the close encounter term will become so insignificant that the resultant conductivity will be practically the same as σ_d (see Fig. 6).

(4) Comparison of Maximum Experimental Conductivity with Theory

The experimental points on Fig. 6(a), (b), and (c) were evaluated from the oscillograms as described in Sec. VI, and they represent the maximum conductivity attained by the gas behind the incident shock wave.

To begin our discussion of the results, let us first examine the 1-cm series shown in Fig. 6(a). It is seen that at high temperatures (solid points) where the oscillograms showed that equilibrium conductivity was reached, the experimental results agreed completely with theory within the limit of accuracy of the theory. At lower temperatures (open points) it is seen that the experimental values fell below the theoretical curve. This discrepancy can be readily explained by the fact that the gas did not have time to reach thermal equilibrium in the region bounded by the shock front and the contact discontinuity as indicated by the typical oscillogram of Fig. 5(a) or (b). At still lower temperatures, it is seen that the experimental conductivity apparently agreed with the theoretical value again. It is believed that this unexpected agreement was due to the effect of impurities which increased the electron concentration by virtue of their lower ionization potentials.

In Fig. 6(b) are shown only a few experimental points for the 0.1-cm series. These experimental points also showed good agreement with theory in the high temperature range. The point at 8800°K showed similar discrepancy from theory as in the 1-cm series at the same temperature.

For the 10-cm series, as shown in Fig. 6(c), the high temperature range (above 10 000°K) was not accessible because of structural limitations of the apparatus. In the accessible range, the experimental points indicated that the departure from thermal equilibrium was less than that in the 1-cm series.

In 1938, Mohler[17] had measured the resistivity of highly ionized cesium vapor in high current density discharge tubes. The results of his measurements at pressures from 0.0045 to 0.080-mm Hg (i.e., Fig. 1 of reference 17) are also plotted in Fig. 6(b) for comparison with our results at high degrees of ionization. In spite of the large difference in operating pressure (notice that the pressure p_2 behind the shock wave is of the order of 100 times the undisturbed pressure p_1, so that our pressure range was actually higher than Mohler's by a factor of 10^3 or more) for the two cases, the comparison is still possible because of the fact that the dependence of conductivity of a highly ionized gas on density is weak [Eq. (8)]. The agreement is better than can be expected because the two sets of results were obtained by completely different approaches and experimental techniques. The reason that the conductivity of highly ionized cesium would agree with that of highly ionized argon has already been mentioned above. The tendency for the cesium conductivity to follow the distant encounter curve (e.g. Spitzer and Härm) to lower temperatures was due to the low ionization potential of cesium (3.87 volts).

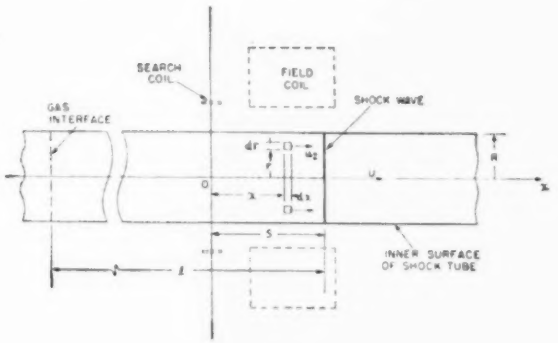


FIG. 8.

APPENDIX I. ANALYSIS OF THE MAGNETIC EXPERIMENT

As depicted in Fig. 8, let x and r be the axial and radial variables of a cylindrical coordinate system whose origin is at the center of the search coil, and whose x axis is directed along the direction of propagation of the shock wave, s be the position of the shock wave at a given time t, R be the inside radius of the shock tube, $U \equiv (ds)/(dt)$ be the speed of the shock wave, u_2 be the flow velocity, and l be the length of the high temperature region bounded by the shock front and the contact discontinuity (see reference 1, Fig. 1).

Now let us consider a small ring of the moving gas at x, with radius r and cross section $dxdr$. If $B_r(x,r)$ is the radial component of the magnetic field due to a unit current in the field coil, and σ is the electrical conductivity of the gas at this point, then the induced electric current in this ring element will be

$$d_i = \sigma u_2 I B_r(x,r) dx dr, \qquad \text{(A-1)}$$

where I is the current applied to the field coil. It should be pointed out that the above expression should include the shielding, or skin effect, of the moving conductor. The magnitude of this effect can be estimated from the familiar solution of a conductor carrying alternating current.[18] By taking $w = U/b$ as the frequency of the alternating current [where b is the effective width of the magnetic field as defined in Eq. (3)], it is found that the current density distribution will be distorted only a few percent due to the skin effect, even for the strongest shock waves in our experiments ($\sigma \approx 80$ mhos/cm, $U \approx 5.5 \times 10^5$ cm/sec). In spite of this indicated small

[17] F. L. Mohler, Bur. Standards J. Research **21**, 873 (1938).

[18] G. Joos, *Theoretical Physics* (Hafner Publishing Company, New York, 1934), Chap. XVI, Sec. 5.

skin effect, the metallic rod experiment (see Sec. III, and also the later part of this Appendix) was performed in the same σU range as that of the strong shock waves to improve the accuracy of the calibration.

The current element di induces a magnetic field the intensity of which can be calculated according to the Biot-Savart law. Thus, if we denote the magnetic flux passing through the search coil circuit due to a unit current ring of radius r and distance x from the search coil by $\gamma(x,r)$, the total induced magnetic flux passing through the search coil when the shock wave is at s will be

$$\phi(s)=\int\gamma(x,r)di$$
$$=u_2I\int_{s-l}^{s}\int_0^R\gamma(x,r)\sigma B_r(x,r)drdx. \quad \text{(A–2)}$$

In writing the above expression, it has been assumed that the flow velocity is uniform and that there is no transverse motion in the gas. As a first approximation the higher order magneto-hydrodynamic interactions, such as the retardation of the flow by the magnetic field, and the change of field current I by the induced magnetic field in the gas, etc., have all been neglected. Approximate calculations showed that these interactions affect the results only by 2 percent or less even for the strongest shocks.

For shock waves of constant strength (i.e., $U=$constant, or very nearly constant, when the shock wave passes through the magnetic field), the flow configuration relative to the shock front is steady (up to the contact discontinuity) so that the thermodynamic state of the gas is a function of the distance from the shock front $\xi\equiv s-x$ only. Therefore, the electrical conductivity σ in Eq. (A–2) may be expressed as $\sigma=\sigma(\xi)=\sigma(s-x)$, when transverse nonuniformity is neglected. Taking $\sigma(s-x)$ out of the integral with respect to r, and writing

$$g(x)\equiv\int_0^R\gamma(x,r)B_r(x,r)dr, \quad \text{(A–3)}$$

Eq. (A–2) becomes

$$\phi(s)=u_2I\int_{s-l}^{s}\sigma(s-x)g(x)dx. \quad \text{(A–4)}$$

According to Faraday's law, the induced emf in the search coil circuit is given by

$$V=A\frac{d\phi}{dt}=AU\frac{d\phi}{ds}, \quad \text{(A–5)}$$

where $U\equiv ds/dt$, and A is a constant proportional to the number of turns of the search coil. From Eqs. (A–4) and (A–5), we obtain

$$V(s)=AUu_2I\left\{\sigma(0)g(s)-\sigma(l)g(s-l)+\int_{s-l}^{s}\sigma'(s-x)g(x)dx\right\}. \quad \text{(A–6)}$$

To evaluate the response function $g(x)$, one may measure $B_r(x,r)$ and $\gamma(x,r)$ and then evaluate the integral numerically according to Eq. (A–3). It has been found, however, that the function $g(x)$ can be obtained much more directly and easily by measuring the induced emf in the search coil when a metallic (nonmagnetic) rod of known electrical conductivity and of approximately the same radius as the shock tube is shot through the magnetic field at known speed. Figure 3 shows a typical oscillogram of such a measurement. This voltage signal was produced by an aluminum alloy rod of finite length (18.9 cm) so that the rear end of the rod produced an opposite but otherwise identical signal when the rod left the magnetic field. When such a rod is sufficiently long (compared with the effective width of the magnetic field), the two voltage pulses will be sufficiently separated so that either of them may be used to evaluate $g(x)$. Under this condition, if u_c denotes the speed of the rod (which in this case is identical with the speed of advance of the conductivity front U_c), σ_c the electrical conductivity of the rod material, and I_c the applied magnetic field current during this measurement, then according to Eq. (A–6), the voltage pulse (either the front pulse or the rear pulse, but not both) is directly proportional to the response function, thus

$$V_c(s)=Au_c^2I_c\sigma_cg(s), \quad \text{(A–7)}$$

where s is now the position of the end of the rod. Substituting Eq. (A–7) into Eq. (A–6) we obtain

$$V(s)=\frac{Uu_2I}{u_c^2I_c\sigma_c}\left\{\sigma(0)V_c(s)-\sigma(l)V_c(s-l)+\int_{s-l}^{s}\sigma'(s-x)V_c(x)dx\right\}. \quad \text{(A–8)}$$

If we integrate the voltage curve, and denote the integral by $\Phi(s)$, then from Eqs. (A–4), (A–5), and (A–7), we obtain

$$\Phi(s)\equiv\int_0^sV(s)ds$$
$$=\frac{Uu_2I}{u_c^2I_c\sigma_c}\int_{s-l}^{s}\sigma(s-x)V_c(x)dx. \quad \text{(A–9)}$$

By a change of variable $\xi=s-x$, Eq. (A–9) may be

rewritten as

$$\Phi(s)=\frac{Uu_2I}{u_c^2I_c\sigma_c}\int_0^t V_c(s-\xi)\sigma(\xi)d\xi, \qquad \text{(A-10)}$$

which will be recognized as an integral equation of the first kind for the unknown function $\sigma(\xi)$.

Conversely, if the conductivity distribution $\sigma(\xi)$ is known, Eqs. (A-8) and (A-10) can be used to calculate the voltage curve $V(s)$ and the voltage integral $\Phi(s)$. The voltage curves and the functions ψ_1 and ψ_2 shown in Table I for various assumed conductivity distributions were obtained from these equations [ψ_1 and ψ_2 were defined in Eq. (2)].

APPENDIX II. ELECTRICAL CONDUCTIVITY OF A SLIGHTLY IONIZED GAS WITH VARIABLE CROSS SECTION

Consider a slightly ionized gas in a weak uniform electric field **E**. Let **C** be the velocity of an electron immediately after a collision with a neutral molecule at time $t=0$. Due to the external electric field, the electron will be accelerated along the direction of **E** before it encounters another molecule. Thus the velocity of the electron at time t will be

$$\mathbf{v}=\mathbf{c}+\frac{\epsilon\mathbf{E}}{m_e}t. \qquad \text{(A-11)}$$

Taking time average of this velocity over the free-flight period, we get

$$\langle\mathbf{v}\rangle_{\text{time}}=\frac{1}{\tau}\int_0^\tau \mathbf{v}dt=\mathbf{c}+\frac{\epsilon\mathbf{E}}{2m_e}\tau, \qquad \text{(A-12)}$$

where τ is the time interval between two successive collisions. If we denote the mean free path appropriate to the electron speed C by λ_c, the period τ will be given by λ_c/c.

Assuming that the molecules are spherically symmetric, and considering the fact that there will be no appreciable persistence of velocity for the electron after a collision, the average of (A-12) over the direction of **C** will be,

$$\langle\mathbf{v}\rangle_{\text{time, direction}}=\frac{\epsilon\mathbf{E}}{2m_e}\tau=\frac{\epsilon\mathbf{E}}{2m_e}\frac{\lambda_c}{c}. \qquad \text{(A-13)}$$

If we further take an average of the above quantity over the electron speed distribution, we obtain the drift velocity for the electrons, and hence the current density in the gas is given by

$$\begin{aligned}\mathbf{j}&=n_e\epsilon\int_0^\infty \langle\mathbf{v}\rangle_{\text{time, direction}}f(c)dc\\&=\frac{n_e\epsilon^2\mathbf{E}}{2m_e}\int_0^\infty\frac{\lambda_c}{c}f(c)dc,\end{aligned} \qquad \text{(A-14)}$$

where $f(c)$ is the electron speed distribution function. Since the electrical conductivity σ is defined by $\mathbf{j}\equiv\sigma\mathbf{E}$, Eq. (A-14) yields

$$\sigma=\frac{n_e\epsilon^2}{2m_e}\int_0^\infty\frac{\lambda_c}{c}f(c)dc. \qquad \text{(A-15)}$$

Now if the collision cross section of the neutral molecule for scattering electrons is $Q(c)$ which is a function of the electron speed c, the mean free path λ_c will be given by $\lambda_c=1/n_0Q(c)$, so that

$$\sigma=\frac{\alpha\epsilon^2}{2m_e}\int_0^\infty\frac{1}{cQ(c)}f(c)dc, \qquad \text{(A-16)}$$

where $\alpha\equiv n_e/n_0$ is the degree of ionization. It is seen that when the cross section Q is a constant, Eq. (A-16) reduces to (for a Maxwellian velocity distribution),

$$\sigma=\frac{1}{(2\pi)^{\frac{1}{2}}}\frac{\alpha\epsilon^2}{(m_ekT)^{\frac{1}{2}}}\frac{1}{Q}, \qquad \text{(A-17)}$$

which differs from Chapman and Cowling's solution [Eq. (6)] only by a factor of $\frac{3}{4}$ as a result of our free-path approximation. Multiplying Eq. (A-16) by the correction factor of 4/3[19] and comparing it with Eq. (6), it is seen that the effective cross section for the gas molecule is given by

$$Q=\left(\frac{2m_e}{\pi kT}\right)^{\frac{1}{2}}\left\{\int_0^\infty\frac{1}{cQ(c)}f(c)dc\right\}^{-1}. \qquad \text{(A-18)}$$

For Maxwellian electron speed distribution, we obtain,

$$Q=\frac{kT}{m_e}\left\{\int_0^\infty\frac{c}{Q(c)}\exp-\left(\frac{m_ec^2}{2kT}\right)dc\right\}^{-1}. \qquad \text{(A-19)}$$

[19] Note that for nonrigid sphere molecules, this correction factor would be appreciably different from 4/3 (e.g. 2 for inverse fifth power, and 4 for Coulomb interaction). Since the exact force law for electron-argon atom interaction is not known, the hard sphere correction of 4/3 was used to compute the dotted curve in Fig. 7 from Normand's results. It should be pointed out, however, that this correction factor does not concern the solid curve in Fig. 7, and hence it does not affect the conductivity curves in Fig. 6.

Theory and Probe Measurements in a Magnetic Ion Source

MAX HOYAUX, ROBERT LEMAITRE, AND PAUL GANS
Nucleonics Research Laboratories, A.C.E.C. Charleroi, Belgium
(Received October 23, 1953)

The "superstate" in a magnetic ion source has been investigated experimentally and theoretically. The space-charge density is approximately constant throughout the source. The radial plasma density does not differ significantly from a Gaussian distribution. The radial distribution of space potential is approximately parabolic. The theoretical results have been tested by probe measurements, and good agreement has been found.

I. THEORY

1. Radial Motion of Electrons

THE normal state and the superstate have been investigated under the assumption that the former is a pure electronic discharge and the latter is a bipolar discharge with nearly equal electronic and ionic densities.

The discharge volume is limited by a cylindrical anode of radius R and length L, and two flat potential walls at $z=\pm L/2$. The electrons are assumed fed in only along the axis of the anode. The magnetic field of induction B is constant and parallel to the axis. As is shown in Fig. 1, the experimental arrangement is as near as possible to the above assumptions.

The motion of electrons in vacuum in a cylindrical chamber under the action of a constant magnetic field and their own space charge has been investigated in some detail.[1,2] A special solution has been found[1] where the electronic mass statistically rotates around the axis at the Larmor speed $-qB/2m$, and where the space-charge density is equal to $\delta_0=\epsilon_0 qB^2/2m$. Under such conditions, the magnetic field force exactly balances the sum of the electrical and centrifugal forces, where the latter two are equal to each other. ϵ_0 is the permittivity in vacuum; q is the charge of the particles; m is the mass of the particles. Mks are used throughout, except for gas pressures.

The corresponding behavior in a gas has been investigated under the assumption that the mean free time is independent of the electron speed. The statistical distribution of the mean free paths has been neglected as usual.

In the immediate neighborhood of the axis, the electrons have not yet undergone any collisions and behave as if in a vacuum. Thus the space charge density in the central region is equal to δ_0. This central region, or "core," extends up to some radius r_0 where the number of electrons inside the cylinder of radius r_0, and length L is equal to that emitted along the axis during the time τ. If an electron collides with an atom, then in general the electron looses angular momentum. The balance of forces is possible again only if the distance r from the axis increases. A space-charge density δ is thus produced which is below the critical one. As can be seen δ tends toward zero if r increases indefinitely.

Consider an electron describing a path at the mean distance r_i from the axis, where the motion in the z direction is an oscillation at a constant speed from $+L/2$ to $-L/2$, and the motion in the $r\theta$ plane is epicycloidal. If an electron, after colliding with an atom, rebounds isotropically, then the mean distance r_e of the path after the collision can be calculated. The expression obtained is

$$r_e=r_i\left[1-\frac{1}{r_i^2}\left(1-2\int_1^{r_i}\frac{\delta}{\delta_0}rdr\right)\right]^{-\frac{1}{2}}. \tag{1}$$

The mean radial component of the electron speed may be defined as

$$v^-=(r_e-r_i)/\tau. \tag{2}$$

2. Radial Motion of Ions

Of the several methods available to investigate the radial motion of ions the one involving a simple application of the classical diffusion theory seems best.[3,4] The presence of the magnetic field is disregarded.

$$v^+=-\mu^+\left(E+\frac{kT^+}{q}\cdot\frac{dN^+}{N^+dr}\right); \tag{3}$$

μ is the ionic mobility; E, the absolute value of the electrical field; k, the Boltzmann constant; T^+, the ionic temperature; and N^+, the ionic density.

3. Theory of Normal State

If the above first two equations are introduced into the expression for the electron current at radius r and equated to the emitted current I, then on the assumption that the ionic density is negligible there is obtained outside the "core":

$$2\alpha(\delta_0/\delta)-\alpha^2(\delta_0/\delta)^2=1+2\int_1^{r_0}\delta rdr/\delta_0 r_0^2, \tag{4}$$

$$\alpha=I\tau/nL\delta_0 r_0^2. \tag{5}$$

[1] Leon Brillouin, Phys. Rev. **67**, 260 (1945).
[2] Max Hoyaux, Publications de l'a. I. Ms, Mons (1951) Atomics, April and November (1951) and February (1952).
[3] A. Von Engel and M. Steenbeck, Elektrische Gasentladungen Tome I, p. 171 ff.
[4] Max Hoyaux, Rev. gen. elec. **60**, 279, 317 (1951).

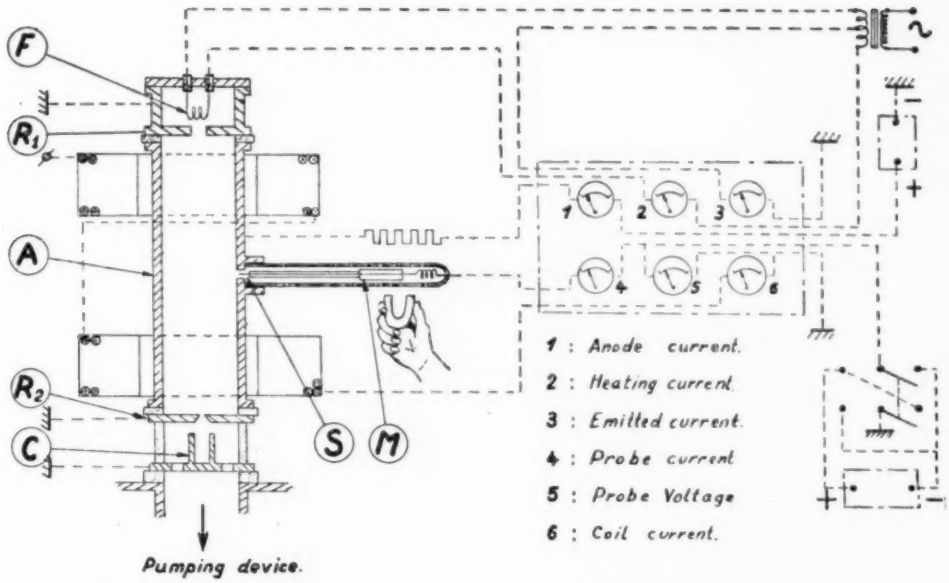


FIG. 1. Schematic diagram of the experimental equipment.

According to the definition of r_0, $\alpha=1$ in the normal state. The solution of Eq. (4) for $\delta=\delta_0$ at $r=r_0$ is $\delta/\delta_0=r_0/r$. Moreover, inside the core, $\delta=\delta_0$ and as no ion density is considered, the electronic density is constant in the core and decreases hyperbolically outside of it.

The number of ions produced per unit time is equal to the number of electrons emitted along the axis per unit time multiplied by the number of ionizing collisions per electron. The latter can be considered approximately equal to the total number of collisions multiplied by some constant probability ω obtained from published data. For example $\omega=\frac{1}{6}$ for hydrogen.[5]

The total number of collisions per electron coming in along the axis in the time τ is equal to the total number of electrons inside the cylinder of radius R and length L divided by the number $I\tau/q$. The total number of ions produced in the source may now be calculated as a current I_p:

$$I_p=\frac{\omega n R^2 L\epsilon_0 q B^2 p}{2m\tau_1} \quad \text{if} \quad r_0\geq R, \tag{6}$$

$$I_p=\omega\left[2RB\left(\frac{nLI\epsilon_0 qp}{2m\tau_1}\right)^{\frac{1}{2}}-I\right] \quad \text{if} \quad r_0<R. \tag{7}$$

The gas pressure is denoted by p, and τ_1 is the value of τ for $p=1$. If $r_0\ll R$, the last term in Eq. (7) becomes negligible. Of course, the actual ionic current is much lower and depends on the efficiency of the extracting device.

4. Theory of Superstate

In the theory of superstate N^+ is assumed approximately equal to N^-. The radial ionic speed v^+ is negligible with respect to the electronic one, and may be considered zero when calculating the radial distribution. If the ion group be considered statistically at rest, the production of ions may be neglected.

Inside the core the assumption $v^+=0$ is introduced in Eq. (3). If E is replaced by the value resulting from $\delta=\delta_0$ and N_0E_0 correspond to the values of N and E for $r=r_0$, then

$$N=N_0\exp[(1-r^2/2r_0^2)qE_0r_0/kT^+]. \tag{8}$$

Outside of the core, the assumption is again made that $v^+=0$ as well as that there is an electronic balance where Eqs. (1) and (2) are to be taken into account and ionizing collisions may be disregarded. The resulting equation is

$$(d/dr)(N/N_0)=[1-2(N/N_0)(r/r_0)^2]/(kT^+/qE_0)(N/N_0)(r/r_0)^3. \tag{9}$$

This equation can be integrated by an iterative method and the solution does not differ significantly from Eq. (8). Furthermore, if by introducing the solution of Eq. (9) into Eq. (3), δ is calculated, it does not differ significantly from δ_0.

The calculation of the production of ions can be performed as in Sec. 3 above. I_p is approximately proportional to I. Previous experimental work seems to have confirmed this conclusion.[6]

II. PROBE MEASUREMENTS

1. Experimental Measurements

A schematic drawing of the equipment is shown in Fig. 1. The filament F emits electrons along the axis, the cylindrical anode A is positive, and the reflectors R_1 and R_2 are negative with respect to the filament. The magnetic field is produced by a pair of Helmholtz coils and a Faraday cylinder C is provided for the

[5] A. Von Engel and M. Steenbeck, Elektrische Gasentladungen Tome I, pp. 30 and 39.

[6] Theodor Finkelstein, Rev. Sci. Instr. **11**, 94 (1940).

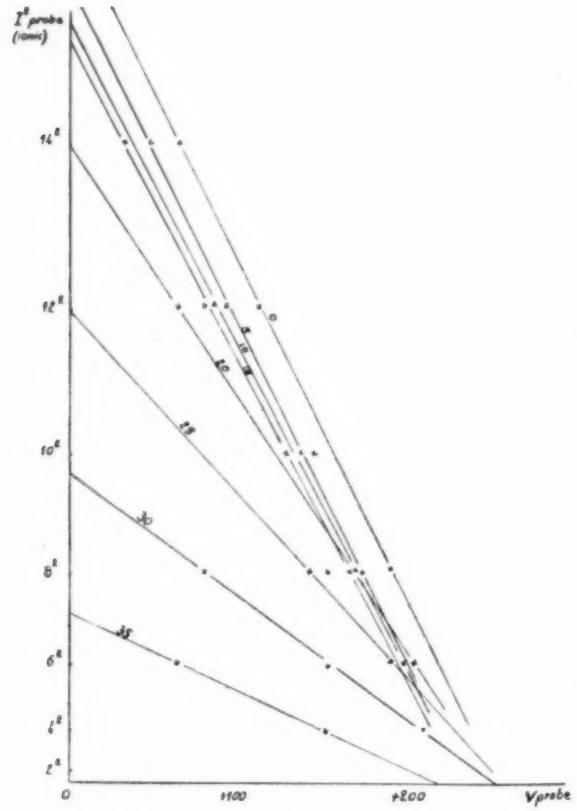


FIG. 2. Plot of the square of ionic probe current *versus* probe-voltage. Coil current 7A. Anode voltage 250 v—anodic current: 500 μA. Emitted current 250 μA. Heating current 0.28A. Cylindrical probe: 0.3×0.3 mm mercury vapor. The values plotted on the curves are the distances from the axis (mm).

extraction of ions. The probe S can travel along a diameter in the median plane of the iron-mass M. The current emitted by the filament and the various voltages on the electrodes are electronically stabilized. Cylindrical and plane probes were both used.

2. Probe Characteristics

According to the theory in part I the motion of the electrons is mainly a rotation around the axis; i.e., for a probe at a definite point, the electrons come mainly from a definite direction at a given speed. According to the theory of Langmuir and Mott-Smith, the electronic characteristic of a cylindrical probe is closely described

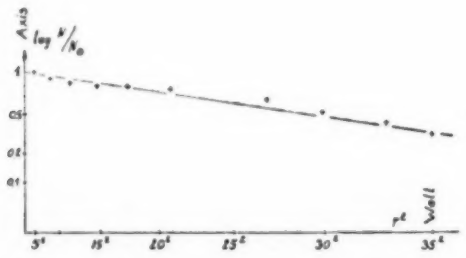


FIG. 3. Radial distribution of the ionic density.

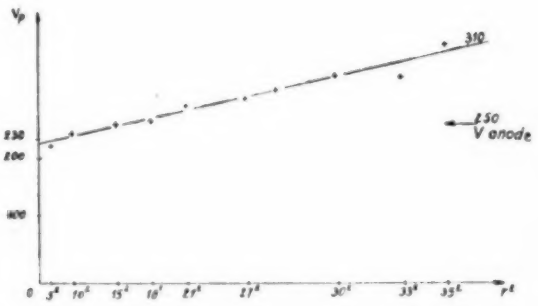


FIG. 4. Radial distribution of the space potential ($+KT^+/q$).

by a parabola with horizontal axis, and that of a plane with guard rings, by a step function.

Because of the random motion of the ions, the ionic branch of the characteristic is similar to that found in ordinary discharges.

In particular, the ionic branch of a cylindrical probe characteristic is such that the square of the probe current plotted as a function of the probe voltage gives a straight line whose slope is proportional to the square of the ionic density and whose intercept on the V axis differs from the space potential by kT^+/q. The results obtained are shown in Fig. 2.

The interpretation of the electronic branch and of the plane probe characteristics with guard rings proved to be less satisfactory and more complicated.

3. Experimental Results

The probe theory is applicable only to the superstate. For practical reasons, the first experiments were performed in mercury vapor.*

According to the theory the ionic, or electronic, density is approximately a Gaussian function of r. A straight line is expected when $\log(N/N_0)$ is plotted as a function of r^2, N_0 corresponding now to the axis. The good experimental agreement is shown in Fig. 3.

Figure 4 shows a plot of the space potential increase in kT^+/q *versus* r^2. If kT^+/q is sufficiently constant a straight line results. The parameter can be calculated by comparing the extrapolated value at the anode with the actual anode potential as read from a voltmeter. The ionic temperature can thus be estimated. Values as high as 720 000°K have been found. Such values are in qualitative agreement with the theory.

ACKNOWLEDGMENT

The kind permission of the Direction of Ateliers de Constructions Electriques de Charleroi to publish the above results is gratefully acknowledged.

The authors would like to take this opportunity of thanking Professor M. L. Pool of Ohio State University for his kind help in reading and correcting the manuscript of this article.

* *Note added in proof.*—Good results have recently been obtained in hydrogen at pressures around 10^{-2} mm Hg. For lower pressures, probe measurements become very difficult.

Experimental and Theoretical Aspects of Shock-Wave Attenuation

IRVINE I. GLASS AND WALTER A. MARTIN
Institute of Aerophysics, University of Toronto, Toronto, Canada

(Received March 3, 1954)

Experimental results are presented of shock wave and contact front velocity measurements, in air, obtained in a 3 in.×3 in. wave interaction tube. A diaphragm pressure ratio range up to 10 000 was employed, while the distance was varied simultaneously from the origin to 142 in. beyond.

It is shown that when shock-wave attenuation occurs, it consists of two portions; (a) a decrement due to formation, and (b) a further attenuation due to the distance traversed by the shock wave.

Concurrently with the attenuation phenomenon, the contact region spreads with time and its front boundary accelerates. The increase in velocity consists of two portions; (a) an increment due to formation, and (b) a further rise in velocity with the distance travelled by the contact front.

A satisfactory empirical relation is developed for the total shock-wave attenuation.

A Rayleigh-type incompressible "pipe"-flow analysis applied to the experimental results overestimates the attenuation for stronger shock waves.

NOTATION

a sound speed
A shock-tube cross-sectional area
C (with arrow over it pointing to the right) contact front moving to the right
d shock-tube diameter or width
m mass flow
M Mach number (u/a)
p static pressure
R_e Reynolds number
S forward facing shock wave (with arrow over it pointing to the right) (particles enter from right)
R backward facing rarefaction wave (with arrow over it pointing to the left) (particles enter from left)
t time
u particle velocity
w shock-wave velocity
x distance along shock tube from diaphragm station
γ specific heat ratio (C_p/C_v)
δ boundary layer thickness
δ^* boundary layer displacement thickness at the contact front
μ viscosity
ν kinematic viscosity (μ/ρ)
ρ density

$P_{41}=\frac{p_4}{p_1}$ diaphragm pressure ratio

$P_{21}=\frac{p_2}{p_1}$ shock pressure ratio

$U_{21}=\frac{u_2}{a_1}$ particle velocity

$W_{11}=\frac{w_1}{a_1}$ shock-wave Mach number

Notes

1. The relations governing the Rankine-Hugoniot curves (R-H curves) are given by Eqs. (1.65) and (1.73) of reference 1 for the shock wave and Eqs. (1.63) and (1.65) for the contact front.

2. Where two or more R-H curves appear on a figure they are displaced from their common origin for clarity.

1. INTRODUCTION

A COMPREHENSIVE study of the shock tube is given in reference 1. An (x,t)-plane diagram of an ideal flow appears on Fig. 1, while a typical (x,t) plane schlieren photograph of a real flow is shown in Fig. 2.

In the ideal case when the diaphragm is removed the shock wave is assumed plane instantly, the rarefaction wave is of the centered type and a contact surface

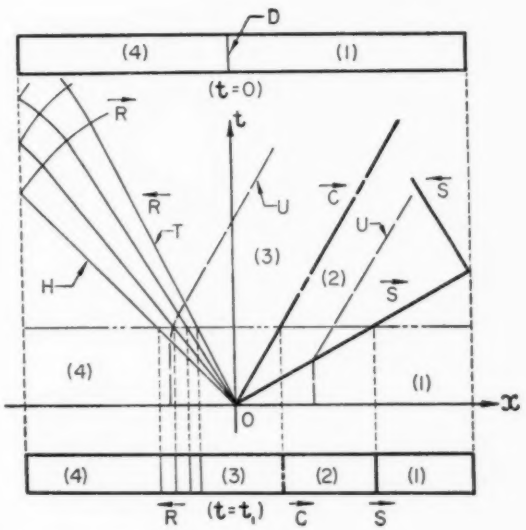
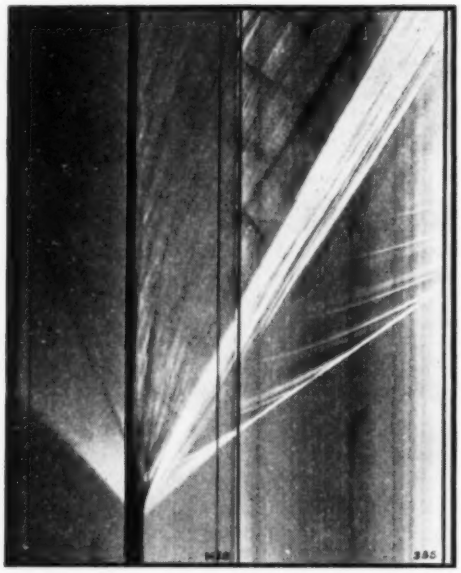
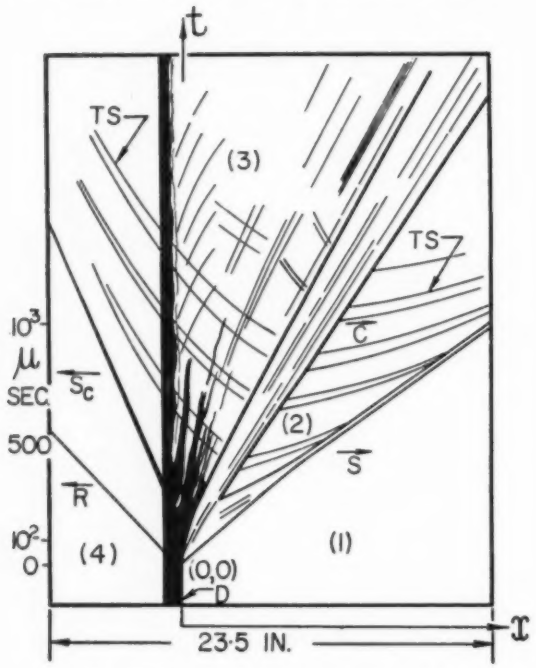


FIG. 1. The wave system produced by bursting a diaphragm in a shock tube. (4)=Chamber. (1)=Channel. D=Diaphragm. S (with arrow over it pointing to the right)=Shock wave. C (with arrow over it pointing to the right)=Contact front. U=Particle path. R (with arrow over it pointing to the left)=Rarefaction wave. H=Head. T=Tail.



(a) Composite photograph of two schlieren records.



(b) Explanatory diagram for Fig. 2(a).

FIG. 2. The wave system produced in a real shock tube from the instant the diaphragm ruptures.

$p_4 = 1500$ mm Hg. $p_1 = 300$ mm Hg. $P_{41} = 5.00$.
$a_{4,1} = 1132$ fps. $W_{11} = 1.39$. $U_{21} = 0.60$.
$P_{21} = 2.1$. $M_2 = 0.50$. $M_3 = 0.65$.
TS = Transverse shock waves.
D = Cellophane diaphragm.
R = Rarefaction wave. C = Contact front.
S = Shock wave. Sc = Condensation shock wave.
(1) = Channel (air). (4) = Chamber (air).

separates the two uniform states. The shock wave propagates with undiminished velocity down the channel of infinite length.

In a real flow curved shock fronts overtake and coalesce to form the primary shock wave. The formation process is shown quite clearly in Fig. 2. Although, the formation region resembles a net of characteristic lines in a compression wave, the head of the wave is already supersonic. Even for weaker waves accelerations produced at the origin are of the order of 7×10^7 ft/sec^2 in the first 10 microseconds.

The contact surface spreads and becomes a thick region in a real flow, the rarefaction wave is not centered and the "uniform" states are crossed by transverse waves, which initially form the mechanism for producing a plane shock wave. The cold state between the contact region and the rarefaction wave is filled with striations that are indicative of turbulent flow. Vortices are also present around the diaphragm station. Therefore, the flow in a real shock tube contains viscous processes, and deviations from an inviscid flow theory may be expected.

One of the first quantities to be determined by shock tube investigators was the shock-wave velocity, since it was relatively easy to measure. A knowledge of the initial conditions and the shock-wave velocity makes it possible to calculate all the physical quantities of the flow in a shock tube.[1] This approach was followed by many experimentalists and they found that the agreement between theory and experiment was quite good for weak waves, but became progressively worse for strong shock waves.

This deviation has been referred to somewhat loosely as shock-wave "attenuation." As a result of the present set of experiments it is possible to define the term "attenuation" in a more rigorous manner. It has been found that associated with each weak or strong shock wave is a characteristic length (x_f) over which the shock forms, accelerates, and achieves a maximum velocity.[1] (Based on the present results a weak shock wave will be defined as one for which $P_{21} < 3$.) The formation distance (x_f) was found to be longer for strong shocks.

For weak shock waves the achieved maximum velocity is the one predicted theoretically; that is, a given diaphragm pressure ratio (P_{41}) produces the theoretical shock-wave velocity (W_{11}).

For strong shock waves ($P_{21} > 3$), it is noticed that although the shock wave does achieve a maximum speed at the end of its formation distance, it falls short of the theoretical wave speed corresponding to the given diaphragm pressure ratio. Instead, the shock wave accelerates to a maximum wave speed, and beyond the formation length it decelerates and the shock velocity decreases monotonically with distance as sketched in Fig. 3(b).

Recapitulating, it may be concluded that for strong shock waves the over-all defect or total attenuation in shock speed consists of two portions: (a) a decrement due to formation, and (b) a further attenuation due to

[1] Glass, Martin, and Patterson, "A Theoretical and Experimental Study of the Shock Tube," Institute of Aerophysics, University of Toronto, UTIA Report No. 2 (1953).

the distance traversed by the shock wave. These will be referred to as a formation decrement and distance attenuation, respectively.

2. SOME FACTORS CONTRIBUTING TO TOTAL SHOCK-WAVE ATTENUATION

It is reasonable to assume in the case of strong shock waves (especially where the kinematic viscosity behind the shock wave is high) that the formation decrement is due to viscous action during the shock formation process and the distance attenuation results from a boundary layer buildup. The possible growth of the boundary layer in a shock tube is illustrated on Fig. 4. It is seen that at the position of the shock wave the boundary layer has zero thickness, since the gas particles in the tube have not been affected yet. At the

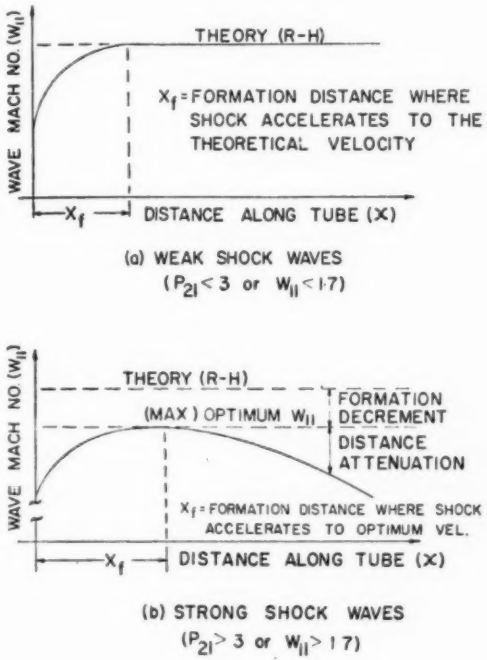

WAVE MACH NO. (W_{11})
THEORY (R-H)
X_f = FORMATION DISTANCE WHERE SHOCK ACCELERATES TO THE THEORETICAL VELOCITY
X_f
DISTANCE ALONG TUBE (X)
(a) WEAK SHOCK WAVES
($P_{21} < 3$ or $W_{11} < 1.7$)
WAVE MACH NO. (W_{11})
THEORY (R-H)
(MAX.) OPTIMUM W_{11}
FORMATION DECREMENT
DISTANCE ATTENUATION
X_f = FORMATION DISTANCE WHERE SHOCK ACCELERATES TO OPTIMUM VEL.
X_f
DISTANCE ALONG TUBE (X)
(b) STRONG SHOCK WAVES
($P_{21} > 3$ or $W_{11} > 1.7$)

FIG. 3. Total shock-wave attenuation in a shock tube.

contact front the boundary layer thickness (δ) is a maximum, since there the particles have been in motion with respect to the wall for the total length of time (t_1) under consideration. Just behind the contact front the Reynolds number increases discontinuously, and for an ideal flow without turbulence in region (3), it may be safely assumed that the boundary layer thickness will decrease in state (3) and through the rarefaction wave, until it is zero again at the head of the wave.

Thus the boundary layer in a shock tube at a given time (t) is a function of position (x). The thickening of the boundary layer with time will diminish the induced mass flow behind the shock wave and as a result its speed will be reduced to balance the lower mass flow. No definite mechanism for reducing the shock speed is postulated. However, since shock decay is caused by a

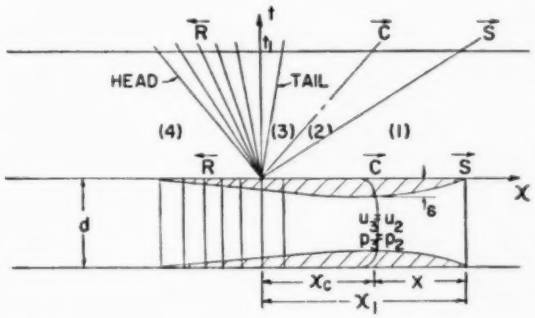


FIG. 4. Boundary layer growth in a shock tube at $t = t_1$.

rarefaction wave, it is generally considered that the contact front slows up. The deceleration produces a rarefaction wave analogous to the classical motion of a piston in a cylinder.[2] It will be shown subsequently that the very opposite happens in a real flow; that is, the contact front actually accelerates. For this reason, the above concept appears rather artificial. Nevertheless, even if the mechanism for slowing up the shock wave is not known, it is reasonable to state that the boundary layer growth reduces the mass flow, which in turn decreases the shock speed and is the cause of the distance attenuation.

The present problem may be compared with Rayleigh's analysis of an infinite flat plate which has been started instantaneously from rest to a uniform velocity u_i at $t=0$.[3] In the present analysis the equivalent problem is considered; that is, the plate is at rest and a flow is generated instantaneously such that its speed is zero at $t=0$, and $u=u_i$ for $t \geq 0$. This is illustrated for a cylindrical shock tube on Fig. 5, and was solved by Donaldson and Sullivan.[4] It is seen that at small times the boundary layer is thin and uniform over the infinite tube. For large times the boundary layer approaches the center of the tube and the motion becomes a pipe flow. In a real shock tube the boundary layer does not grow uniformly and it occupies only a finite length of the tube. Furthermore, the end-boundary conditions are complex, due to the presence of the shock wave and the contact region. However, it will be shown subse-

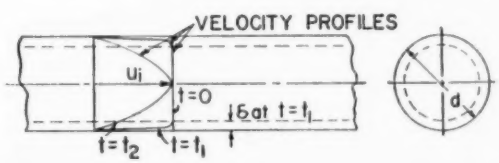


FIG. 5. Boundary layer profiles with increasing time for the equivalent Rayleigh problem in a cylindrical tube.

[2] R. Courant and K. O. Friedrichs, *Supersonic Flow and Shock Waves* (Interscience Publishers, Inc., New York, 1948).

[3] S. Goldstein, *Modern Developments in Fluid Dynamics* (Oxford University Press, New York, 1938) Vol. 1, pp. 181–186.

[4] C. du P. Donaldson and R. D. Sullivan, "The Effect of Wall Friction on Strength on Shock Waves in Tubes and Hydraulic Jumps in Channels," Natl. Advisory Comm. Aeronaut. Tech. Note 1942 (1949).

quently that some useful results may be obtained by considering the two problems as equivalent.

A dimensional analysis of an instantaneously started smooth tube indicates that the boundary layer (δ) (see Fig. 5), neglecting heat transfer, is a function of the following quantities:

$$\delta = f(p,\rho,\mu,\gamma,u,d,l), \tag{1}$$

or

$$\delta/d = g(M,R_e,\nu t/d^2,\gamma). \tag{2}$$

Therefore, if tube roughness and heat transfer are neglected, then for a given gas (γ), the distance attenuation which depends on the boundary layer may be expected to vary with the Mach number (M), the Reynolds number (R_e), and the time parameter ($\nu t/d^2$).

In the case of the shock tube, the gas particles at the contact front are in motion with respect to the wall for the greatest length of time, and this is the time which would correspond to the time considered in the equivalent Rayleigh problem. If the total shock-wave attenuation is neglected this time (t_1) is computed as (Fig. 4)

$$t_1 = \frac{x_c}{u_2} = \frac{x_1}{w_1} = \frac{X}{w_1 - u_2}. \tag{3}$$

It is convenient to measure t_1 in terms of the shock position from the diaphragm ($t_1 = x_1/w_1$), and the time parameter may be expressed as

$$\frac{\nu_2 t}{d^2} \equiv \frac{\nu_2 x}{Aw}, \tag{4}$$

where A is the cross-sectional area proportional to d^2.

The foregoing type of analysis cannot be applied directly to the formation decrement, since it depends on the diaphragm rupturing process, which is of a complex nature and introduces three-dimensional viscous effects.

3. EXPERIMENTAL PROCEDURE

In order to check some of the factors contributing to the formation decrement and the distance attenuation, a series of experiments were performed. These experiments included a study of the effect of various types and combinations of diaphragms, and the influence of extraneous disturbances such as the shredded diaphragm and the diaphragm rupturing device on the formation decrement. To study the distance attenuation, shock velocities were measured with a wave speed camera and schlieren system from the origin to 142 in. beyond the diaphragm station (0 to 50 tube widths) over the range $1 < P_{41} \leq 10\,000$.

A detailed description of the wave speed camera and the schlieren system utilized in conjunction with the present tube is given by Glass.[5] The 12-in. diameter camera drum rotates at about 4200 rpm and carries a 70 mm × 3½ ft strip of film externally. The film speed is known to better than one-tenth of one percent. Angles were measured to 0.5 of a minute or better by means of a Hilger T500 measuring projector. This made it possible to measure large wave speeds even at $P_{41} = 10\,000$ to better than two percent. The greatest possible error occurred in measuring low-channel pressures (0.40 mm Hg minimum). An error of nine percent was possible, but this was alleviated by taking several runs and utilizing average values. A comprehensive discussion of the experimental methods may be found in reference 1.

4. EXPERIMENTAL RESULTS

The experimental results may be summarized as follows. Cellophane diaphragms MSTL 600, used singly or in combination, gave the closest approximation to an ideal diaphragm removal when they were loaded to about half or greater of their static rupturing strength. The energy contributions to the flow from the breaking cellophane diaphragms may be neglected. The introduction of discontinuities at the origin on small protrusions in the flow have no measurable effect on the shock-wave velocity. The formation distance where the shock becomes plane and achieves a straight path on an (x,t)-plane schlieren record varies. For weak waves ($P_{21} < 3$), about 4 tube widths are required; for strong shock waves (where $P_{21} \approx 17$), the formation distance extends over about 8 tube widths.

The effects of distance attenuation and the formation decrement for $1 < P_{41} \leq 10\,000$ and 33 in. $\leq x \leq$ 142 in. are summarized on Fig. 6. It is seen that for weak waves the formation decrement and distance attenuation are absent. For strong waves the formation decrement and distance attenuation increases with P_{41} and x. The dashed curve shows the wave speeds that would be produced as a result of variable specific heat. It accounts for only a small portion of the experimental attenuation even at very high P_{41}. Thus Fig. 6 confirms the graphic representation of the total shock-wave attenuation shown on Fig. 3.

During the present experiments it was necessary to employ a low channel pressure ($p_1 = 0.40$ mm Hg) in order to achieve a diaphragm pressure ratio, $P_{41} = 10\,000$. A low value of p_1, gives a correspondingly high kinematic viscosity (ν_2) behind the shock wave. The dimensional analysis showed that ν_2 and x were two very important quantities which would affect the distance attenuation and similarly ν_2 would influence the formation decrement.

The same value of p_1 was utilized for a given P_{41} over the entire distance (x) range. Therefore, for any given P_{41} the only variable was x. It was possible to develop a linear empirical relation between P_{41}, W_{11} and x as shown on Fig. 7, for the present experiments. The Rankine-Hugoniot (R-H) curve may be approximated by

$$\log P_{41} = 1.34(W_{11} - 1) + 0.176. \tag{5}$$

[5] I. I. Glass, "The Design of a Wave Interaction Tube," Institute of Aerophysics, University of Toronto, UTIA Report No. 6 (1950).

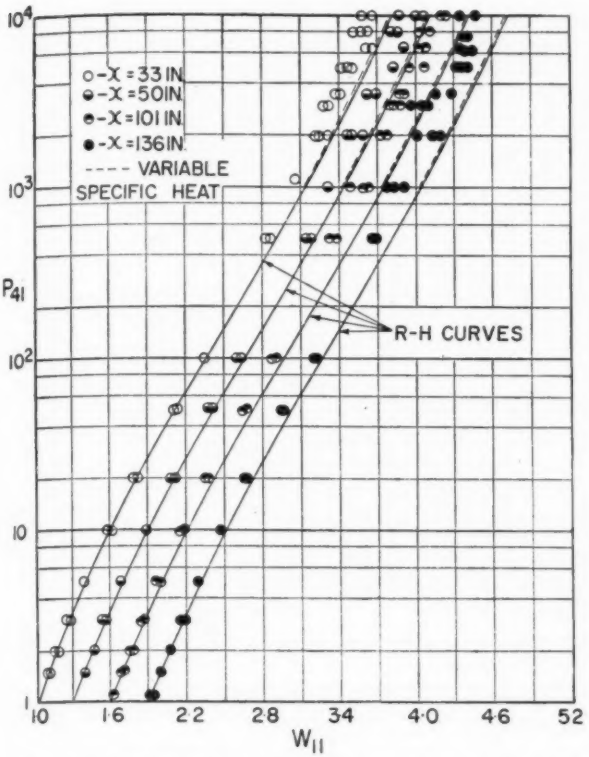


FIG. 6. Variation of initial shock wave velocity (W_{11}) with diaphragm pressure ratio (P_{41}), $x=33$, 50, 101, and 136 inches. (In air.)

The experimental results are described quite well by the relation

$$\log P_{41}=1.33x^{0.03}(W_{11}-1)+0.08. \tag{6}$$

When this equation was applied to the results of reference 10 at a distance of 10 feet from the diaphragm station it was found that the agreement was quite good. It would seem from this that this formula could be used to determine the distance attenuation for tubes of similar cross-sectional area when the value of $\nu_2 t$ was also similar.

5. INCOMPRESSIBLE VISCOUS-FLOW ANALYSIS

Although Eq. (6) describes adequately the total attenuation for the present experiments, it was desirable to find a general solution which would be applicable to any shock tube. The first attempt in this direction was to solve the incompressible, viscous equations of motion for an impulsively generated flow in a square tube, with the following initial and boundary conditions:

$$\frac{\partial u}{\partial t}=\nu\left(\frac{\partial^2 u}{\partial z^2}+\frac{\partial^2 u}{\partial y^2}\right), \tag{7}$$

$$u(z,y,0)=u_i \quad \text{for} \quad 0<z<z_0; \quad 0<y<y_0,$$

$$\left.\begin{matrix}u(0,y,t)\\ u(z,0,t)\\ u(z_0,y,t)\\ u(z,y_0,t)\end{matrix}\right\}=0 \quad \text{for} \quad t>0.$$

The solutions for the velocity distribution and the mass flow are analogous to those in heat conduction theory[6] and are as follows:

$$\frac{u}{u_i}=\sum_{a=1}^{\infty}\sum_{b=1}^{\infty}\frac{4}{ab\pi^2}(1-\cos a\pi)(1-\cos b\pi)$$

$$\times\exp\left[-\frac{\nu t}{A}(a^2+b^2)\pi^2\right]\sin\frac{a\pi z}{z_0}\sin\frac{b\pi y}{y_0}, \tag{8}$$

$$\frac{m}{m_i}=\frac{4}{\pi^4}\sum_{a=1}^{\infty}\sum_{b=1}^{\infty}\frac{(1-\cos a\pi)^2(1-\cos b\pi)^2}{a^2b^2}$$

$$\times\exp\left[-\frac{\nu t}{A}(a^2+b^2)\pi^2\right]. \tag{9}$$

If the distance attenuation in a shock tube is due to a defect in mass flow behind the shock wave, then this may be expressed, for air, as[1,4]

$$\frac{m}{m_i}=\frac{P_{21}-1}{P_{21_i}-1}\times\frac{P_{21_i}+6}{P_{21}+6}\left(\frac{6P_{21}+1}{6P_{21_i}+1}\right)^{\frac{1}{2}}. \tag{10}$$

By considering the impulsive flow and the shock tube flow as equivalent, it was possible to plot Fig. 8. The results indicate that the viscous flow analysis (Eq. (9)) gives a good description of the flow up to diaphragm pressure ratios of 200.

Thus the viscous analysis follows the experimental results in the range of P_{41} where $M_2\leq 1.2$. However, in this range the experimental values of $\nu_2 t/A$ were also very small (about 10^{-4}). Therefore, $m\approx m_i$, and the distance attenuation would be negligible. Hence it should not be misconstrued that agreement is good in the incompressible range only, but that the total attenuation is small for weak waves.

For $P_{41}>200$ the predicted attenuation is too high

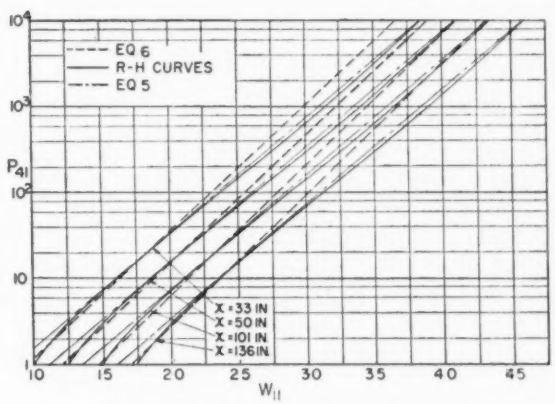


FIG. 7. Variation of initial shock-wave velocity (W_{11}) with diaphragm pressure ratio (P_{41}) for $x=33$, 50, 101, and 136 inches. (In air.)

[6] R. V. Churchill, *Fourier Series and Boundary Value Problems* (McGraw-Hill Book Company, Inc., New York, 1941).

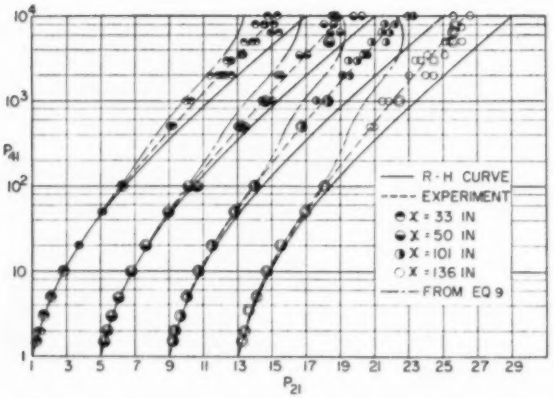


FIG. 8. Variation of initial shock pressure ratio (P_{21}) with diaphragm pressure ratio (P_{41}) for $x=33$, 50, 101, and 136 inches. (In air.)

and confirms the previous remarks, that when the time parameter ($\nu_2 t/A$) is large, the viscous effects may be expected to be excessive. In this case, the flow becomes a nonstationary pipe flow analogous to the stationary Poiseuille flow.

Although this was a distance attenuation analysis, it is noted that the results for $x=33$ in. are already far too great, and here the main effect is really due to the formation decrement. A more realistic approach is to assume that the flow consists of an inviscid core and a viscous boundary layer. The boundary layer thickness is a function of x at a given t between the shock wave and the contact front. However, it is assumed that the displacement thickness (δ^*) at the contact front controls the defect in mass flow, which in turn causes the shock-wave attenuation. The mass flow ratio depends on the initial and final strengths of the shock wave and is found from Eq. (10).

The mass defect at the contact front,

$$\Delta m=\rho_2 u_2 \Delta A=\rho_2 u_2 \delta^* D,$$

where D=shock tube perimeter, and ρ_2, u_2, and δ^* are measured at the contact front.

For a square tube,

$$\delta^*=\frac{A}{D}\left(1-\frac{m}{m_i}\right)=R_H\left(1-\frac{m}{m_i}\right)=\frac{y_0}{4}\left(1-\frac{m}{m_i}\right), \quad (11)$$

where R_H=the hydraulic radius.

The above definition of δ^* may be questioned. However, it can be tested experimentally by interferometric means, and some work in this direction is presently underway at the Institute of Aerophysics.

By utilizing the experimental results of Fig. 8 in conjunction with Eqs. (10) and (11), it was possible to plot a family of curves of δ^* *vs* x for varying shock strength or flow Mach number (M_2), as shown on Fig. 9.

It is seen that at the higher values of M_2, a linear relation exists between δ^* and x or $\nu_2 t/A$. As M_2 decreases the linearity grows poorer, and below $M=1.29$, the scatter was too large to warrant a plot. This occurs when the experimental curves approach the (R-H) curves (Fig. 8), making it difficult to measure accurately the distance between them. For the Mach number range ($1.3<M_2<1.5$) plotted on Fig. 9,

$$\frac{d\delta^*}{dx}=0.0004.$$

The equivalent value for stationary flow is about tenfold.[7,8]

If the lines in Fig. 9 are extended towards the origin, it is seen that a large portion of δ^* is attributable to the formation decrement. This decrement increases with M_2 or ν_2 as expected. The growth $d\delta^*/dx$ exhibits a sensibly constant value, although an increase with ν_2 would have been more reasonable from physical considerations. Nevertheless, Fig. 9 confirms that for stronger shock waves the total attenuation consists of a formation decrement and a distance attenuation.

Further investigations are being conducted on the effects of varying ν_2 at a given M_2 over the distance range, and of varying M_2 at a given ν_2 with distance. The results will provide additional necessary experimental data for Fig. 9.

A computation of δ^* from the experimental results of reference 10 are also plotted on Fig. 9 and they exhibit reasonably good agreement with the present analysis. This confirms the previous remark that for tubes in which ν_2 and A are similar, the shock-tube results may be expected to agree.

In summary, it may be stated that the weakness of the Rayleigh-type, incompressible, impulsively started flow analyses lies in equating the mass flow from the

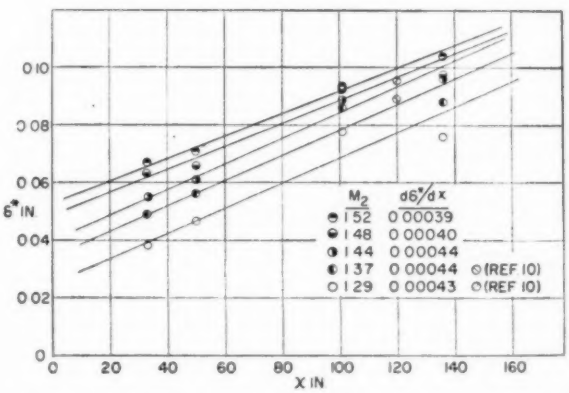


FIG. 9. Variation of boundary layer displacement thickness (δ^*) with distance (x) from the diaphragm for values of M_2. (In air.)

[7] J. Ruptash, "Boundary Layer Measurements in the UTIA 5 inch×7 inch Supersonic Wind Tunnel," Institute of Aerophysics, University of Toronto, UTIA Report No. 16 (1952).

[8] G. C. Goldbaum, "Comparison of Theoretical and Experimental Performance of a Variable Geometry Diffuser at $M=5.0$," University of Texas, Report DRL-331 (1953).

shock tube with the mass flow in an infinite tube. In the latter the boundary layer grows uniformly. In the shock tube the boundary layer growth over the finite length is nonuniform, and furthermore the end-boundary conditions are very complex. Very little gain can be expected from a consideration of the compressible, Rayleigh-type analysis, since the compressible boundary layer displacement thickness for a flat plate is even greater.

Other measurements of shock-wave attenuation and approaches to the shock-tube boundary layer problem have been made.[9–13] There is no theory in existence today which can unify the results obtained by the different investigators. The assumptions in the theoretical treatments generally neglect some of the important boundary conditions and ignore the variation of the physical quantities that occur in a real flow.[1,14,15] Consequently, the agreement between theory and experiment differs for each observer and depends on the initial conditions, the range of diaphragm pressure ratios, shock-tube geometry, and lengths that were utilized during the experiments.

6. CONTACT FRONT VELOCITY MEASUREMENTS

It is shown in reference 14, that the transition through a contact front involves the solution of a set of nonlinear differential equations in heat conduction. The solution indicates that the contact front grows in time due to the molecular transport of heat. Consequently, the advance of the front of the contact region is more rapid than the rear boundary.

In a real flow [Fig. 2], the contact front from the burst diaphragm is very thick and irregular. Nevertheless, the discontinuity in the density derivative $(\partial\rho/\partial x)$ at the boundaries of the contact region nearly always appear on a schlieren record of the (x,t) plane. As the diaphragm pressure ratio increases, the rear boundary merges with the rough flow of state (3). Finally, at low channel pressures ($p_1<1$ mm Hg) even the front boundary is difficult to distinguish. Typical results for the speed of the front boundary of the contact region (U_{21}), (in an ideal flow U_{21} is also the particle speed) as a function of P_{41} and x are shown on Fig. 10. Briefly, the behavior is opposite to that of the shock front. For weak and strong contact regions there is a deceleration of the initial jet outflow following the rupture of the diaphragm.[1] At the end of the formation distance an optimum velocity is reached. For weak contact fronts the optimum velocity is the one predicted by the shock-tube theory (R-H curves), regardless of distance (x). For strong contact fronts, the optimum velocity is greater than the theoretical value, and the velocity further increases with distance.

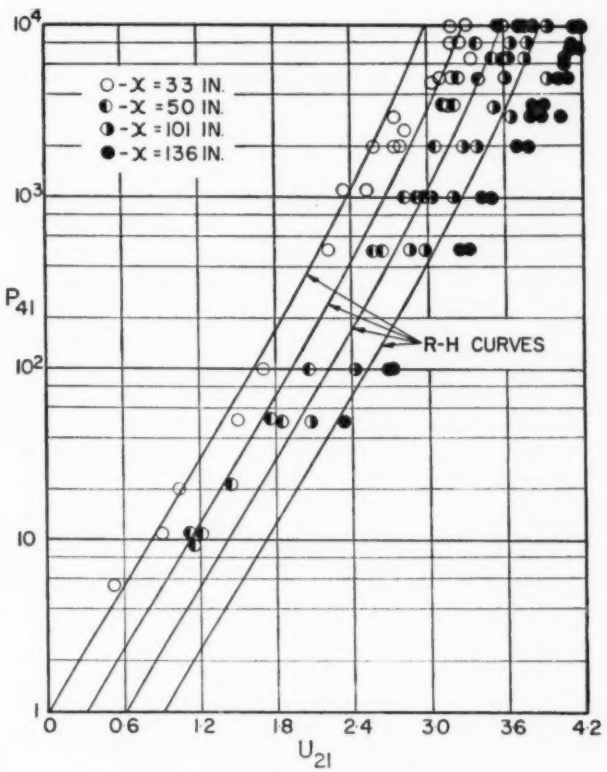


FIG. 10. Variation of initial particle velocity (U_{21}) with diaphragm pressure ratio (P_{41}) for $x=33$, 50, 101, and 136 inches. (In air.)

The velocity increase of the front boundary of the contact region cannot be explained on the basis of the heat conduction solution for a "laminar" contact front, since its order of magnitude is small.[14] In a real flow the contact region is turbulent and its growth with time cannot be predicted at present.*

The spreading of the contact region will not only invalidate the concept of uniform-flow regions in a shock tube but will materially affect the shock-wave velocity and the transition profiles through the rarefaction wave.[1,15]

Before a realistic mathematical model can be formulated for shock-wave attenuation, contact front acceleration, and variable flow properties it will be necessary to determine experimentally the pressure, density, temperature, and wave-velocity profiles between the shock wave and the rarefaction wave. A good start in

[9] R. K. Lobb, Phys. Rev. **84**, 612 (1951).
[10] R. J. Emrich and C. W. Curtis, J. Appl. Phys. **24**, 360 (1953).
[11] Huber, McFarland, and Levine, "Investigation of the Attenuation of Plane Shock Waves Moving Over Very Rough Surfaces," Natl. Advisory Comm. Aeronaut., R.M. SL53D13a (1953).
[12] R. E. Duff, J. Appl. Phys. **23**, 1373 (1952).
[13] R. N. Hollyer, Jr., "A Study of Attenuation in the Shock Tube," University of Michigan, Engineering Research Institute Report M720-4 (1953).
[14] J. G. Hall, "The Transition Through a Contact Region," Institute of Aerophysics, University of Toronto, UTIA Report No. 26 (1954).
[15] J. E. Mack, "Density Measurement in Shock-Tube Flow with the Chrono-Interferometer," Lehigh University, Institute of Research Tech. Report No. 4 (1954).

* Because of the no-slip condition at the tube walls, the contact front in a real flow extends back to the diaphragm station and adds to the complexity of the shock tube problem.

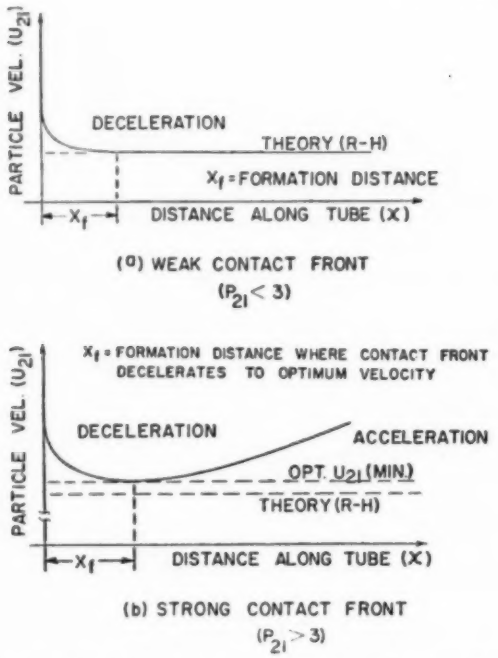


FIG. 11. The behavior of the contact front in a shock tube.

this direction has already been made in references 1, 15 and 16.

7. CONCLUSIONS

The present wave-speed experiments have shown that for real flows two very noticeable phenomena occur in a shock tube simultaneously, namely, shock-wave attenuation and contact front acceleration.

For weak shock waves (wave Mach No. <1.7) both are absent and the wave-speed results for the shock wave and the contact front agree with the simple shock-tube theory.

For strong shock waves (wave Mach No. >1.7), the above noted effects increase monotonically with Mach number (kinematic viscosity) and the distance traversed by the shock wave and the contact front. Total shock-wave attenuation consists of a formation decrement plus a further defect in wave speed due to the distance travelled by the shock which is called distance attenuation. The maximum speed attained by the shock wave after acceleration over the formation distance is always less than the theoretical value, and the speed then progressively decreases with distance. On the other hand, the front boundary of the contact region decelerates to an optimum velocity at the end of the formation distance which is greater than the theoretical value, and the speed then progressively increases with distance.

It is shown in reference 1 that the train of transverse shock waves which follows the primary shock wave regardless of distance (0- to 50-tube widths) also contributes to the total wave attenuation. For strong shock waves, the transverse waves, shock attenuation, and contact front acceleration will no doubt affect the uniformity and the magnitude of the flow quantities in the states separated by the contact region.

So far, a unifying theory has not been produced that will account for the various observations of shock-wave attenuation reported by many investigators. A successful theory must give consideration to the flow conditions outlined above that exist in a real shock tube.

Some experimentalists may wonder whether such a precise theoretical treatment would really be warranted, since a calibration curve can usually be obtained for a given shock tube. The answer might be that not only is this an interesting physical phenomenon in itself, but it may have an important bearing on two other associated problems; namely, (a) is it really valid to employ the shock-wave velocity to calculate the magnitude of the flow quantities in the "uniform" states separated by the contact front? (b) are the "uniform" states really uniform? The evidence available of direct measurements of pressure, density, temperature, or flow velocity over the limited diaphragm pressure ratio range does not permit a decision. Until these questions are settled experimentally, the researcher cannot accept these assumptions without reservation, especially for stronger shock waves.

An Observation of Circular Patterns in the Vicinity of Small-Area Alloyed Germanium *p-n* Junctions*

NICK HOLONYAK, JR.†
Electrical Engineering Research Laboratory, University of Illinois, Urbana, Illinois
(Received April 2, 1954)

Alloyed germanium *p-n* junctions of 0.1 to 0.2 mm diameter or less have been alloyed upon crystals of roughly ten times greater linear dimensions such that when these units are suitably electrolytically etched, ring formations (both scattered rings and tracks of overlapping rings) occur on the crystal faces in the neighborhoods of the junctions. Best current evidence indicates that the observed effects are etch produced. An explanation of this is given in terms of the action of bubbles accompanying the etching procedure.

INTRODUCTION

IN most work involving alloyed indium-germanium *p-n* junctions[1] small junction size is not particularly important or often desired. For some experimental purposes, however, small junctions are of interest and give rise, under certain conditions of preparation, to rather striking circular patterns on the surfaces of germanium crystals in the immediate neighborhoods of the junctions. In particular, indium junctions of roughly 0.1 to 0.2 mm radius alloyed upon relatively large specimens of germanium and with, in each case, rather uniform circular patterns on the crystal surface in the region of the junction have been prepared. The patterns are composed of overlapping circles which have a tendency to radiate outward from the junction. In each case the patterns either were made evident or more likely were produced by electrolytic etching, the etching current being introduced through the junction. A more complete description of the patterns is given below following, first, a more detailed description of the junction fabrication process.

FABRICATION OF JUNCTIONS

In all cases in which the circular patterns were observed, the alloyed junctions displaying the effect were constructed as follows:

A specimen of germanium of from 0.5 mm to 2 mm thickness and 0.25 cm^2 to 1 cm^2 area was first chemically etched and then placed on a tin-antimony tinned Kovar heating strip as shown in Fig. 1. A 0.002 inch diameter nickel hairpin with a small bead of indium attached to one end was then brought in contact with the germanium crystal. Dry, oxygen-free hydrogen was introduced about each unit, and finally the entire assembly was heated by passing current through the Kovar base (heating) strip. The magnitude and duration of the heating current were so adjusted that the indium wet the germanium and formed a small alloyed junction.

After alloying, each unit was electrolytically etched. Prior to etching the crystal face to which the junction was alloyed was clear. After etching patterns of the sort displayed by Fig. 2 were obtained. Subsequent chemical etching caused the patterns to fade more or less uniformly throughout their entire extent.

DESCRIPTION OF THE PATTERNS AND THEIR GENERATION

A typical picture of the patterns under discussion is shown in Fig. 2. Figure 2 is a photomicrograph of the crystal surface surrounding an alloyed indium-germanium *p-n* junction prepared as stated above. The junction itself is represented by the dark center region bounded by the partially broken white circular line. The junction diameter, as indicated, is 0.12 mm. The dark region in the immediate vicinity of the junction is undistorted germanium which has been etched to a greater depth than the outlying regions. Beyond the dark center regions occur patterns which appear as overlapping circles or rings having a tendency to be centered on radial lines through the junction. Inward from the main body of patterns there is some tendency for the rings to fade, particularly in the region adjacent to the clear, deeply etched area surrounding the junction.

Figure 3 is a more highly magnified view of the upper left-hand portion of Fig. 2. Figure 3 reveals that the circles of which the patterns are composed have a fine structure which in Fig. 4 is revealed as a series of pits defining the circles. Figure 4 is a view, more highly

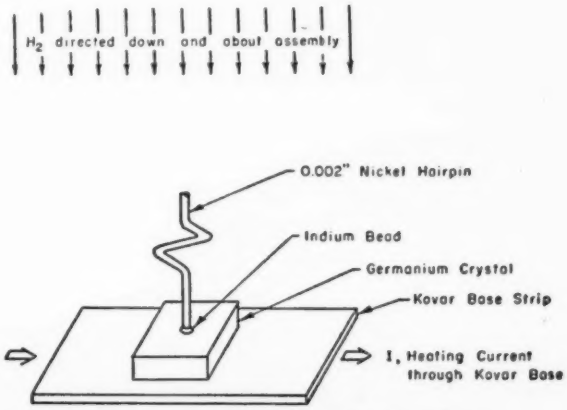


FIG. 1. Assembly for junction fabrication.

* This work was supported in part by U. S. Office of Naval Research funds and in part by a Texas Instruments fellowship.
† Present address, Bell Telephone Laboratories, Murray Hill, New Jersey.
[1] R. N. Hall and W. C. Dunlap, Phys. Rev. **80**, 467 (1950).

FIG. 2. Photomicrograph of alloyed indium-germanium *p-n* junction (0.12 mm diameter) with surrounding circular patterns.

magnified than Fig. 3, of an unidentified region in Fig. 2. Figure 4, at the lower left, reveals an interesting circle of pits which ends abruptly in a somewhat deeper pit. The end etch pit could well be suspected to be of an origin different from that of the main body of pits.

At first it was believed that the occurrence of patterns of the type shown in Fig. 2 was due to structural defects in the germanium, defects introduced during the alloying (heating) operation. It was thought that temperature gradients inherent in a system of the type of Fig. 1 could set up stresses which in turn might give rise to plastic flow and hence to structural defects which would be exposed by etching. Though this explanation seems valid, it suffers in that other workers who have alloyed junctions using setups similar to that of Fig. 1 and who have followed essentially the junction fabrication process outlined above have not observed the effect. Also it is unexplained why electrolytic etching would reveal structural defects and chemical etching would not. This, in fact, would tend to contradict the results of others who have found evidence of dislocations in observations of chemically etched germanium crystals.[2]

A more likely explanation of the effect has recently been suggested by Dr. John Saby. He has suggested that bubble production inherent in rapid electrolytic etching could be responsible for patterns of the type depicted in Fig. 2.‡ In the high current density region surrounding the indium-germanium interface bubbles are generated when sufficiently large etching currents are employed. Presumably bubbles generated at the junction would be kicked out radially and would follow the germanium surface for a short distance. Where a bubble was in contact with the crystal surface no etching would occur; the bubble masks the surface, shields it from the electrolyte and hence from etching. The bubble, being spherical, would cover a circular region on the germanium surface and as it moved would, depending on the motion, leave a path of either poorly or well defined overlapping circles. This model would account very well for the overlapping of the tracks of rings in Fig. 2; a bubble giving rise to a track could very well cover a portion of the path traced by an earlier bubble.

After Dr. Saby's hypothesis had been suggested, further experiments were conducted with closer attention being paid to bubbling during etching. A number of indium-germanium junctions were prepared with no particular attempt to alloy the units under temperature

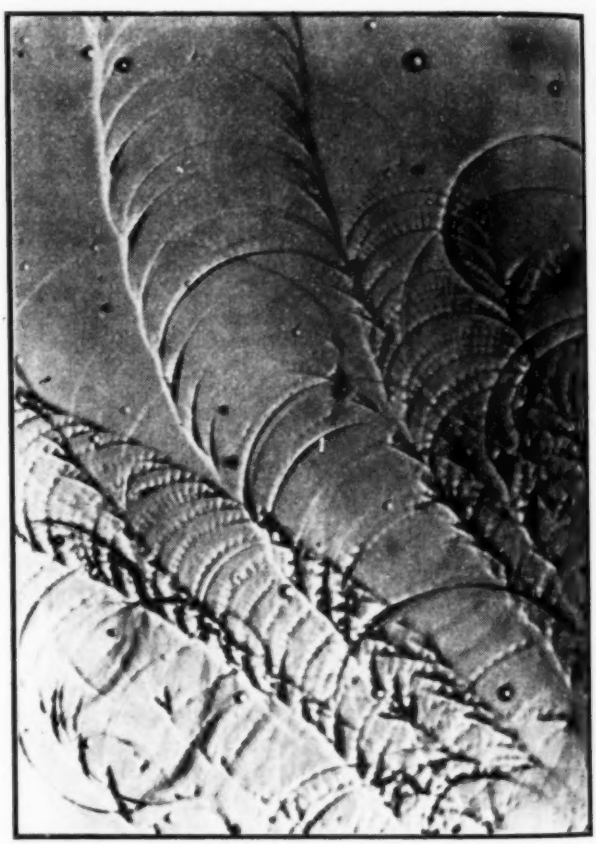


FIG. 3. More highly magnified photomicrograph of region located in upper left of Fig. 2.

[2] Vogel, Pfann, Corey, and Thomas, Phys. Rev. **90**, 489 (1953); and C. J. Gallagher, Phys. Rev. **92**, 846 (1953).

‡ The possibility of etch production of the patterns of Fig. 2 was considered when the effect was first noticed. However, failure to consider the influence of bubbling led to abandonment of etching as an explanation of the cause of the effect.

gradient-free conditions. The units so prepared were observed when etched to appear as follows.

When relatively light etching currents, i.e., those resulting in little or no bubbling, were used, no circular patterns of any sort were observed surrounding the junctions. Introduction of heavier etching currents with accompanying bubble production at the junctions resulted in both scattered circular patterns and tracks of overlapping circles located in close to the junctions. The circular patterns observed did not extend out over the crystal faces to distances as large as those in Fig. 2. This would be expected, however, since the units considered here were not etched as hard as were those displaying patterns as shown in Fig. 2. Units of the type shown in Fig. 2 were etched such that vigorous bubbling occurred and were etched until the indium beads were nearly cut away from the germanium crystals. The deep undercut below the indium bead in a unit might in some way cause bubbles to be ejected along the crystal surface to appreciable distances before the bubble buoyancy carried it upward. It was noticed in the units possessing the circular patterns in close to the junction that during etching some bubbles were lodged at the base of the indium beads and slowly expanded and finally floated free. An expanding bubble could account also for tracks of overlapping rings. In such case, in order to account for equal sized rings in a track, it would be necessary to assume that increasing buoyancy in an expanding bubble is such as to keep roughly the same sized region of the crystal masked by a bubble.

It was noticed, finally, that in the units having no rings when first etched and then possessing scattered rings and tracks of rings after heavier etching it was possible to re-etch with lower currents (no bubbles at the junction) and remove the rings left after heavier etching. This observation would add support to the bubble theory.

It should be mentioned that certain other aspects were studied of the units possessing surfaces of the kind shown in Fig. 2. First of all examination of the rectification characteristics of such units indicated behavior in good agreement with accepted p-n junction theory. Tungsten contacts made to the circular pattern regions of originally n-type crystals displayed rectification characteristics similar to those of conventional tungsten to n-type germanium contacts. X-ray analyses of two units indicated that the circular patterns were located on (100) planes, but this is not likely to be essential to occurrence of the effect inasmuch as the rings occurred in many cases on specimens from different single crystals and of possibly different orientations. Besides occurring in numerous small alloyed indium-germanium units, the effect characterized by Fig. 2 was obtained in several junctions constructed by alloying 0.005 inch diameter gallium-doped gold wire to germanium specimens of the size previously mentioned. The effects in the gold wire-germanium junctions were essentially the same as the effects already described for the indium-germanium units.

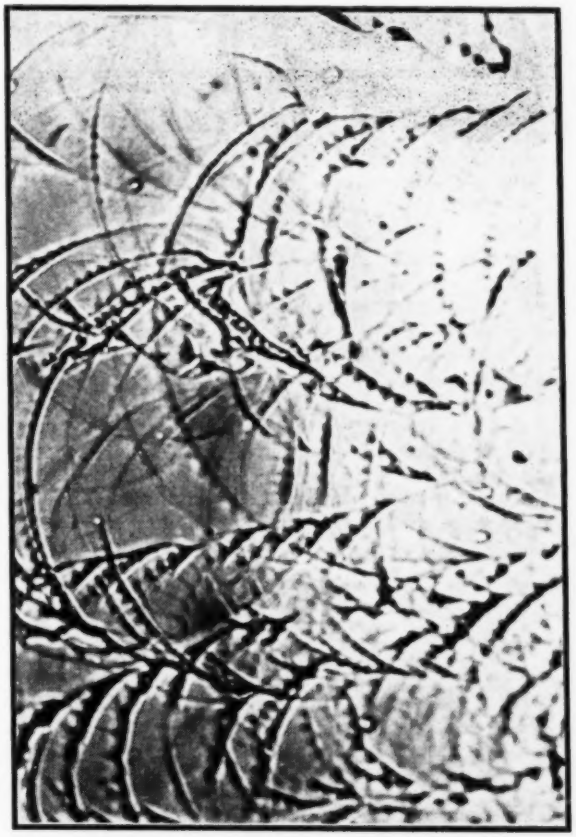


FIG. 4. Highly magnified photomicrograph showing etch pit structure of an unidentified region in Fig. 2.

ACKNOWLEDGMENT

The writer would like to thank those members of the Bell Telephone Laboratories, Texas Instruments, and the General Electric Company who made available the germanium crystals used in this work. He would like also to acknowledge his indebtedness to Dr. John Saby for having provided information regarding alloying of junctions and for helpful discussions on all phases of this work. Finally, he would like to express his thanks to Professor John Bardeen for constant counsel and encouragement.

Letters to the Editor

Time Delay in High-Speed Ferrite Microwave Switches

R. CONWAY LECRAW AND HENRY B. BRUNS
Diamond Ordnance Fuze Laboratories, Washington, D. C.
(Received September 15, 1954)

IN a previous letter[1] an X-band ferrite microwave switch was reported with a switching time of 12 millimicroseconds (mμsec). It was stated that for the particular ferrite[2] there was no apparent time delay of the microwave output pulse with respect to the magnetizing pulse. Since the appearance of this letter, a number of inquiries have been received concerning the magnitude of the time delay which, from physical considerations, is necessarily present. To measure such small delays a method capable of greater time resolution is required. Some results from such a method are outlined below.

Figure 1A shows the current pulse I, from a hydrogen thyratron pulser, in a five-turn primary coil 0.18 in. long and 0.28 in. in diameter. The coil is wound 2.5 in. from the end of a ferrite rod[2]

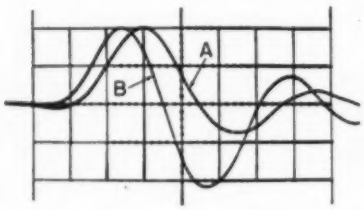


FIG. 1. A: Primary current I (I_{max} = 2.7 amp, H_{max} ≃ 20 oersteds). B: $d\phi/dt$ at a distance 1.5 in. from center of primary. The same external trigger was used for the sweep in A and B. The time scale is 10 mμsec per large division.

0.25 in. in diameter and 9 in. long. Figure 1B shows $d\phi/dt$, the negative of the voltage induced in a one-turn secondary, at a distance s = 1.5 in. from the center of the primary. By noting where $d\phi/dt = 0$, the time delay τ between the flux maximum and the current maximum is seen to be approximately 5 mμsec. In Fig. 2C τ is plotted *vs* s, using a current pulse with the shape of

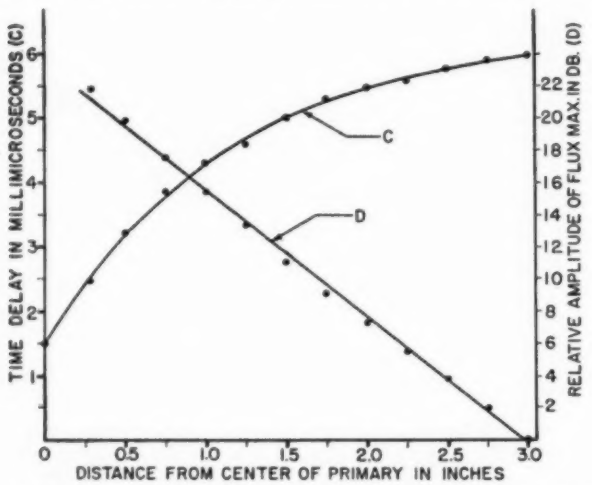


FIG. 2. C: Time delay τ of ϕ_{max} with respect to I_{max}. D: Relative amplitude of ϕ_{max} in db. The time scale used in obtaining Curve C was 5 mμsec per large division or one mμsec per small division.

Fig. 1A and with I_{max} = 2.7 amp, H_{max} ≃ 20 oersteds. Figure 2D shows the relative amplitude of the flux maximum in db as a function of s.

For the above-mentioned microwave switch, H_{max} required for switching (i.e. 90° rotation of the plane of the E vector) is approximately 18 oersteds with a 0.25 in. diameter rod[2] 3 in. long. Hence the section of 9 in. rod from $s = -1.5$ in. to $s = 1.5$ in. can be considered as simulating a rod used in a switch, neglecting end effects. Figure 2C shows that the time delay τ is no greater than 5 mμsec for all points $|s| \leq 1.5$ in. Actually, if this 3-in. section of rod were used in a switch, 90° rotation would occur approximately 3.2 mμsec after I_{max} rather than 5 mμsec. At $\tau = 3.2$ mμsec (corresponds to $s = 0.5$ in.) the magnetization for $|s| < 0.5$ in. is large and decreasing, and for $0.5 < |s| \leq 1.5$ in. it is smaller and increasing. This value of τ is arrived at by a weighting process based on the fact that the magnetization, and hence the rotation per unit length of rod, is much greater near the center than toward the ends. This is seen in Fig. 2D. Thus for a microwave switch using this type ferrite and a current pulse with the magnitude and rise time, 17 mμsec,[3] of Fig. 1A, the switch-off (or switch-on) state occurs approximately 3.2 mμsec after the current maximum.

It is of interest to note from Fig. 2 that both the time delay and the amplitude of the flux maximum[4] are, to good approximations, exponential functions of the distance s. Further experiments have shown this to be true also for larger values of H_{max}.

[1] R. C. LeCraw, J. Appl. Phys. 15, 678 (1954).
[2] A MgMn ferrite, type R-1 (formerly MF 1331), manufactured by General Ceramics Corporation, Keasbey, New Jersey.
[3] This current risetime is the same as used in reference 1 and produces a microwave switching time of 12 mμsec measured from the 1.5 db downpoint of P_0, since for a rotation-type switch $P_0 \sim \cos^2\theta$ and $\theta \sim I$, where θ is the angle of rotation.
[4] The amplitude of ϕ_{max} has also been found to vary exponentially along iron wires. T. R. Lyle, J. M. Baldwin, Phil. Mag. [6] 12, 433–468 (1906).

Kinetics of Magnetization in Some Square Loop Magnetic Tapes

C. P. BEAN AND D. S. RODBELL
General Electric Research Laboratory, Schenectady, New York
(Received September 15, 1954)

THE speed with which a reversal of magnetization occurs by domain wall movement in a ferromagnetic material is determined by the velocity of a domain wall and the number of domain walls involved for a given geometry. The wall velocity is determined by energy dissipation caused by eddy currents[1] and energy irreversibly communicated to the spin system by passage of a domain wall.[2] The number of domain walls is determined by the way in which nucleation takes place in the material.

It has been pointed out previously by Williams, Shockley, and Kittel[1] that for a thick picture-frame single-crystal specimen of silicon-iron, the rate of magnetization is determined in the main by eddy currents, and at low fields one domain wall traverses the width of the specimen. At high fields, a continuous domain wall is considered to move inward from the surface, again limited by eddy currents.

Figure 1 shows the observed rate of magnetization in a constant field of a toroidal core of magnetically annealed 65 Permalloy (65 percent Ni, 35 percent Fe). The particular core represented by these data was composed of 29 wraps of tape, 0.005 in. in thickness and 0.5 in. wide. The two experimental curves plotted show two flux reversals carried out under identical conditions, illustrating the lack of detailed reproducibility in this low driving field region. The solid curve represents the rate of flux reversal based on the Williams, Shockley, and Kittel picture, i.e., instant nucleation of one domain wall at an edge which propagates across the width of the specimen limited only by energy dissipation in the form of eddy currents. The substantial agreement indicates that the domain geometry is quite simple and that this material, because of the anisotropy impressed by magnetic annealing, acts as a magnetic single crystal, although metallurgically it is polycrystalline. Similar agreement is obtained for 0.003 in. and 0.010 in. tapes. This result complements those of Williams and Goertz,[2] who showed magnetically annealed Perminvar (45 percent Ni, 34 percent Fe, 23 percent Co) to be substantially a magnetic single crystal.

An alloy of 50 percent Ni−50 percent Fe can be made into a tape having a rectangular hysteresis loop by appropriate cold

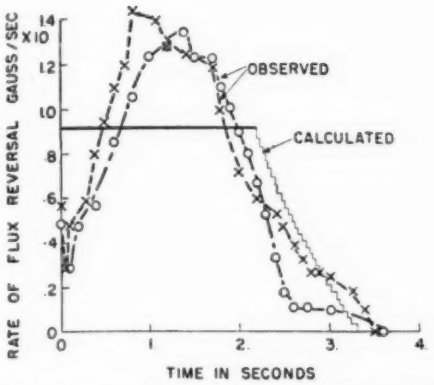


FIG. 1. Rate of change of flux *vs* time. Comparison of calculated and experimental wave shapes for a 0.005 in. magnetically annealed Permalloy tape average excess field of 0.011 oersted.

working and heat treating such that a strong texture is developed with the 100 plane in the plane of the tape and the 100 direction along the length of the tape. A sample of this type did not show the simple behavior previously described. Instead it appears that domains are nucleated at many places within the tape. Confirmatory evidence comes from the observation that the reversal time may be increased as much as a factor of two by placing a large field in the reverse direction prior to application of the driving field. The large field is considered to remove certain nuclei which would otherwise be active. This effect has not been observed in the magnetically annealed 65 Permalloy.

At high driving fields, all of these materials appear to reverse by domain walls moving in from the surface.[2,4] Figure 2 shows the entire range of reversal times in the 0.005-in. tape of 65 Permalloy. The two solid curves represent theoretical curves: (a) instant nucleation on one edge and traversal across the width, and (b) instant nucleation on the surface and soaking toward the center. In each case, the motion is assumed to be limited only by the eddy current. In the high field case the calculated times are some 25 percent shorter than the observed. Part of this discrepancy is caused by a delay in nucleating the domain wall configuration and part by a limitation of its velocity by energy dissipation in the spin system. This will be discussed in an article to be submitted to this journal.

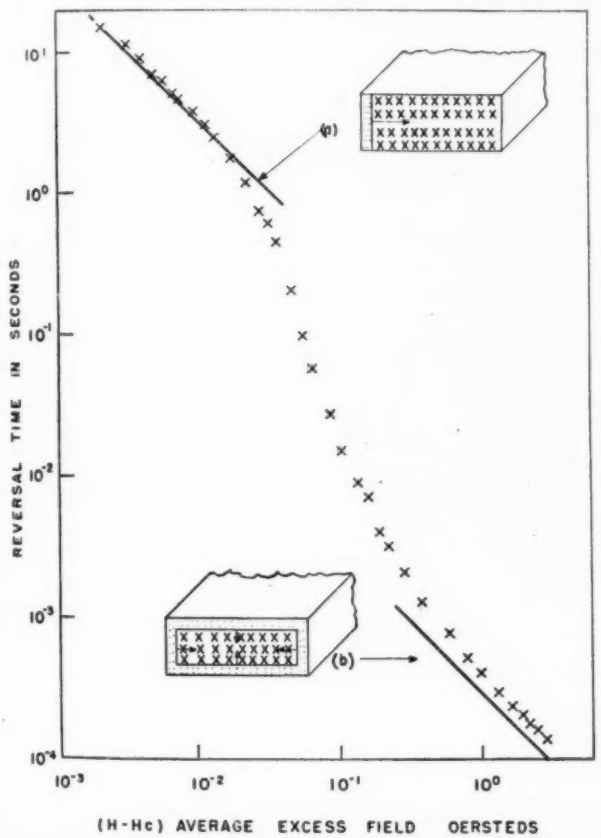


FIG. 2. Total reversal time *vs* average excess field. Characteristic of the 0.005-in. magnetically annealed Permalloy tape with the models that predict the reversal behavior shown.

The authors wish to acknowledge the aid of Mr. G. H. Howe, who prepared the specimens used in this work. In addition we wish to thank Professor H. Eckstein and others of the Armour Institute group for a very pleasant and profitable discussion.

[1] Williams, Shockley, and Kittel, Phys. Rev. 80, 1090 (1950).
[2] L. Landau and E. Lifshitz, Physik. Z. Sowjetunion 8, 153 (1935).
[3] H. J. Williams and M. Goertz, J. Appl. Phys. 23, 316 (1952).
[4] A. G. Ganz, Am. Inst. Elec. Engrs. Trans. 65, 177 (1946).

Shock Waves in Air Produced by Elastic and Plastic Waves in a Plate

WILLIAM A. ALLEN, JOE M. MAPES, AND EARLE B. MAYFIELD
Michelson Laboratory, U. S. Naval Ordnance Test Station, Inyokern, China Lake, California
(Received September 13, 1954)

THIS note reports an experiment that demonstrates existence of a plastic compression wave in certain solids whose dynamic stress-strain curve can be approximated with two straight-line segments. Two shadowgraphs are reproduced in Fig. 1 of shock waves in air produced by elastic and plastic waves

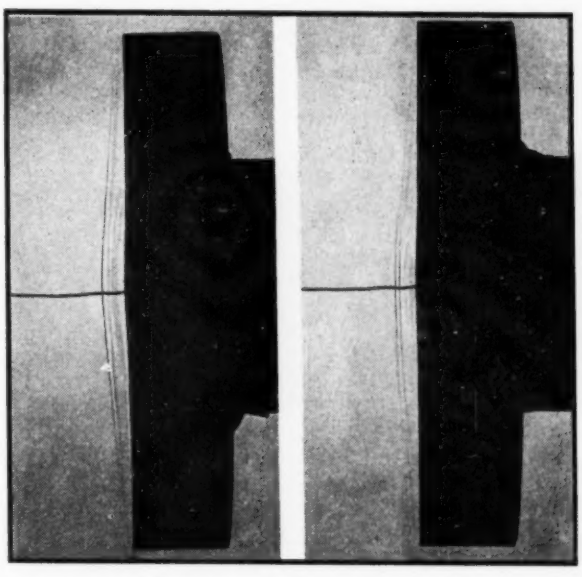


FIG. 1. Air shocks produced by impact of 0.50 cal 2800 ft/sec cylinders on plates. Air pressure: 932.3 millibars. Temperature: 96°F. The horizontal wires were used to trigger the spark when the plate moved forward to establish contact. (a) Copper plate. (b) Brass plate.

in a plate. A 0.1 μsec point light source,[1] developed for this kind of application, was used to stop the motion produced by impact of a 0.50 cal cylindrical projectile at ordnance velocity into the back surface of the plate.

The theory of plastic wave propagation in long thin bars[2–4] was developed in 1941 and later extended[5] to lateral distances large compared to that traveled by the waves. The theoretical solution

can be expressed in terms of strain waves whose velocity of propagation is given by the relation

$$c=\left(\frac{1}{\rho}\frac{\partial\sigma}{\partial\epsilon}\right)^{\frac{1}{2}},$$

where ρ is the initial density of the material, and σ and ϵ are the longitudinal dynamic stress and strain, respectively. The dynamic stress-strain relation is usually inferred from measurements of the static stress-strain curve in tension. The relation in compression is similar to that in simple tension.[6] The static stress-strain curve of copper has a continuous first derivative, while that of brass can be approximated by two straight-line segments.[7] If the same behavior occurs under dynamic conditions, it is plausible that a well-defined plastic wave can be produced in brass, but not in copper.

Experimental evidence[8,9] has been published of a plastic wave that appears at the free surface of a 2-in. steel plate about 4 μsec later than arrival of the initial elastic wave. The loading in these previous experiments was produced by a 2-in. cylindrical charge detonated in contact with the plate. In most cases the metal spalled near the free surface. The present technique generates a high-intensity step impulse that allows the metal to flow, but not to fracture. A 0.50 cal cylindrical steel projectile 1 in. in length was fired into the back surface of 1 in. thick, 5.5 in. diam plates to produce shock waves in air from the free surfaces. Figure 1a illustrates that a single well-defined shock wave is formed from a copper plate. When a cylindrical projectile is fired into a brass plate, Fig. 1b, two shock waves are formed. The first wave, associated with the initial elastic wave in the metal, is followed after an interval of about 3 μsec by a second wave that corresponds to the plastic wave. The second wave is associated with the actual plastic deformation of the free surface. The difference in behavior of the two metals is roughly that expected from theory.

The technique has been applied successfully to other metals such as steel and lead. Steel produces two shock waves; lead produces a single shock. This result is in agreement with previous work[10] which indicates that the shock wave set up by high velocity impact has an initial velocity in steel which is less, and an initial velocity in lead which is greater, than the velocity of plane elastic waves.

[1] Frank N. Miller, Phys. Rev. **93**, 945 (1954).
[2] T. von Karman and P. Duwez, J. Appl. Phys. **21**, 987 (1950).
[3] G. I. Taylor, J. Inst. Civil Engrs. (London) **26**, 486 (1946).
[4] K. A. Rakhmatulin, Appl. Math. and Mech. (Moscow) **9**, No. 1 (1945).
[5] D. S. Wood, J. Appl. Mech. **74**, 521 (1952).
[6] A. Nadi, *Plasticity* (McGraw-Hill Book Company, Inc., New York, 1931), first edition, p. 124.
[7] M. O. Withey and James Aston, *Johnson's Materials of Construction* (John Wiley and Sons, Inc., New York, 1947), eighth edition, p. 742.
[8] William A. Allen and Clyde L. McCrary, Rev. Sci. Instr. **24**, 165 (1953).
[9] William A. Allen, J. Appl. Phys. **24**, 1180 (1953).
[10] Pack, Evans, and James, Proc. Phys. Soc. (London) **60**, 1 (1948).

Flash Associated with High-Velocity Impact on Aluminum*

W. W. Atkins
Applied Physics Laboratory, The Johns Hopkins University, Silver Spring, Maryland
(Received August 21, 1954)

By the use of a Libessart spark[1] and suitable optical system, a time history was recorded of the growth of the luminous impact flash produced by high-velocity sphere perforation of an aluminum target.

The high intensity of the impact flash required a special optical system[2] (Fig. 1) to prevent complete blackening of the photographic film which occurred when using ordinary shadowgraph techniques. Although the spark unit and film holder provide only a single shadowgram during one perforation, a continuous history of the phenomena was obtained by recording a number of impacts with varied time delay after contact by the projectile. Zero time occurred upon initiation of the impact flash which was recorded by a photoelectric cell and oscillograph. The current fluctuation occurring in the spark discharge circuit was used to induce a second timing mark on the intensity-time oscillogram. The time elapsing between the flash initiation and spark discharge could be read directly on the oscillogram.

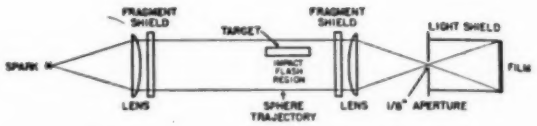


FIG. 1. Schematic diagram of optical system.

Figure 2(a) is an untimed picture taken at an angle of 20° from the normal to the target. It shows the impact flash to be approximately annular.

The remaining photographs of Fig. 2 are special shadowgrams taken parallel to the target plate and normal to the trajectory with varied time separations. The projectile is a 0.500-in.-diameter steel sphere having a velocity of approximately 5000 ft/sec fired against a 0.500-in.-thick 24S-T4 Al target. Each photograph shows the parallel component of light from the impact flash which passes through the aperture; the striae of the lens system are also shown. A groove in the target-mounting holder makes it possible to see a small region of the exit side of the target. In Figs. 2(b), (c), and (d) the spalling visible on the exit surface is of no significance as this spall is caused by a previous round fired into the same target. Small dark spots visible in photographs (b), (c), (d), (e), and (f) are caused by pits in the Lucite shields produced by fragments from previous rounds.

Figure 2(b) is a shadowgram taken at an estimated time of 1 μsec before sphere impingement. Figure 2(c) was taken when the projectile had penetrated approximately $\frac{1}{8}$ in. and at 2 μsec after the instant of contact. The impact flash has propagated to a diameter of about $1\frac{1}{4}$ in. and thickness of about $\frac{1}{4}$ in. Figures 2(d) and 2(e) show the degree of penetration after 4 μsec and 14 μsec, respectively. Also from Fig. 2(e) shock waves caused by target

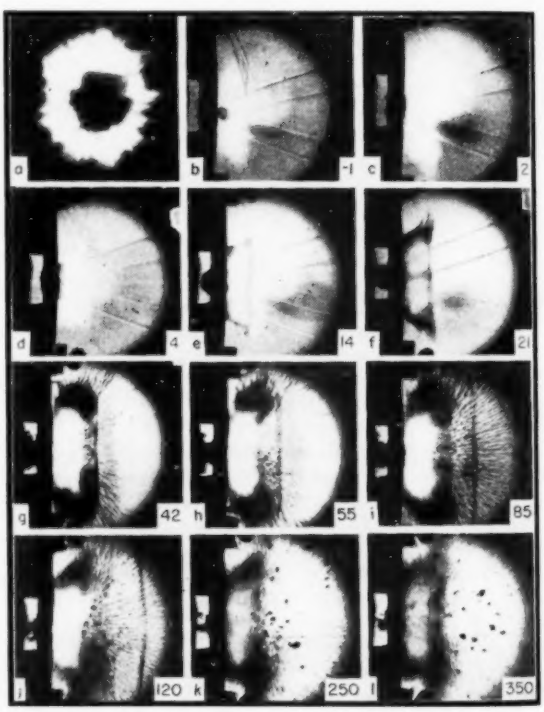


FIG. 2. Growth of an impact flash. The time interval between the moment of sphere contact and the spark discharge is recorded in microseconds and stated on each shadowgram.

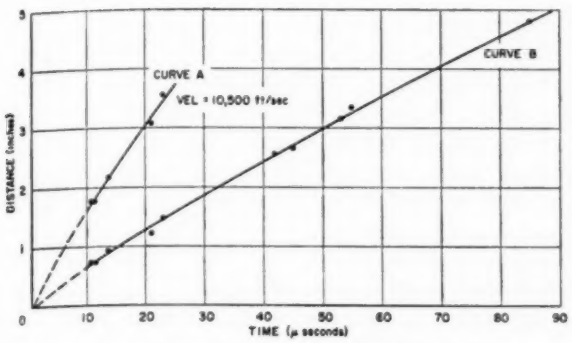


FIG. 3. Displacement-time curves of small fragments.

vibrations are visible and small fragments from the target are emerging through the flash front at an angle of approximately 60°–70° from the trajectory. At a time of 21 μsec [Fig. 2(f)] it is not certain that the sphere has perforated the target although a shock wave from the back bulge is evident. Figure 2(g) again shows this shock as well as a shock wave caused by the flash propagation on the entrance side of the target. Detachment and propagation of the front shock are recorded by Figs. 2(h), (i), and (j). Large fragments from the entrance side of the target emerge through the flash and are clearly observable after 250 μsec. The duration of the luminous flash is estimated at 4 to 5 milliseconds; however, some disturbance in the flash region as well as fragements were observed as late as 22 milliseconds after impact.

Since it has been observed that most of the small fragments diverge radially at an angle of 60° to 70° from the trajectory of the projectile, a displacement-time plot (Fig. 3) for these fragments during the early stage of flash development is given by Curve *A*. The fragment velocity appears to be about 10 500 ft/sec. along this path. By a displacement-time plot of the fragments which appear normal to the target (Curve *B*) and projection of this component of velocity to an angle of 60°–70° a resultant velocity is obtained which is roughly the same as that of Curve *A*.

* Research program conducted at the Naval Research Laboratory, Washington, D. C., and jointly supported by the Bureau of Ordnance, U. S. Navy, JHU/APL under contract NOrd-7386 and NRL.

[1] Melton, Prescott, and Gayhart, The Johns Hopkins University, Applied Physics Laboratory, Bumblebee Report No. 90, December, 1948.

[2] Private communication from R. B. Edmonson, The Johns Hopkins University, Applied Physics Laboratory, Silver Spring, Maryland, January, 1954.

Effect of Asymmetric Oxygen Scattering on the Slowing-Down Length of Neutrons in Water

H. C. VOLKIN
Lewis Flight Propulsion Laboratory National Advisory Committee for Aeronautics, Cleveland, Ohio

(Received July 8, 1954)

THE average square of the slowing-down length to 1.44 ev (indium resonance) in water for neutrons produced by thermal fission of U_{235} has been determined experimentally by Hill, Roberts, and Fitch[1] and turns out to be about 20 percent larger than the value for this quantity as calculated from Fermi's theoretical expression for the slowing-down length, L_s, of neutrons from a monoenergetic source in hydrogeneous media.[2] Fermi's expression for L_s was deduced under the following assumptions: (1) the hydrogen scattering is spherically symmetric in the center-of-mass system; (2) the nonhydrogen nuclei have infinite mass; (3) the nonhydrogen scattering is spherically symmetric; and (4) all collisions are elastic ones. Marshak[3] has derived the same expression for L_s starting from the appropriate transport equation.

In a previous communication[1] L_s was calculated with restriction (4) relaxed so as to take into account the effect of neutron capture. In the present communication an expression for L_s will be derived in which the deviation from spherical symmetry of the elastic scattering by nonhydrogen nuclei, as well as the effect of capture, is taken into account; i.e. assumption (3) is eliminated and assumption (4) is weakened to admit capture collisions. For this purpose it is necessary to consider the transport equation with the scattering function F given by the following expression, in which the notation follows closely that used by Marshak,[3]

$$F(\mu;\, u,u') = (1/2\pi)c(u')e^{-(u-u')}\delta[\mu - e^{-\frac{1}{2}(u-u')}] + c_0(u')\delta(u-u')\sigma(\mu,u')/\sigma_0(u'). \quad (1)$$

Here u equal $\log(E_i/E)$ and u' equals $\log(E_i/E')$, where E_i is the source energy and E' and E are the neutron energies before and after a collision, respectively; μ is the cosine of the scattering angle in the laboratory system; c and c_0 are the probabilities that a collision is an elastic one involving a hydrogen and a nonhydrogen nucleus, respectively; and σ and σ_0 are the differential and the total scattering cross sections, respectively, for the latter process. If the total, the hydrogen scattering, and the nonhydrogen scattering mean free paths are l_t, l, and l_0, respectively, then $c=l_t/l$ and $c_0=l_t/l_0$. Following the procedure of Marshak[3] one can derive from the resulting transport equation the following expression for L_s^2 referred to a final neutron energy $E_f=E_ie^{-u}$:

$$L_s^2(u) = \frac{1}{3}\left\{\frac{l^2(0)}{[1-b_1(0)][1-b_0(0)]} + H(u) + \int_0^u du' \frac{c(u')}{1-b_0(u')}H(u')\right\}, \quad (2)$$

where, with $\sigma_n(u)$ designating the nth Legendre moment of σ,

$$b_n(u) = c_0(u)\sigma_n(u)/\sigma_0(u),$$

$$H(u) = \frac{l^2(u)}{[1-b_1(u)][1-b_0(u)]} + \frac{l(0)l(u)}{[1-b_1(0)][1-b_1(u)]} \times \exp[-G(u)] + \frac{l(u)}{[1-b_1(u)]}\exp[-G(u)] \times \int_0^u du' \frac{l(u')c(u')}{[1-b_1(u')][1-b_0(u')]}\exp[G(u')],$$

$$G(u) = \int_0^u du'\left[\frac{1}{2} - \frac{c(u')}{1-b_1(u')} + \frac{c(u')}{1-b_0(u')}\right].$$

In computing the slowing-down length in water from Eq. (2) the required cross sections for hydrogen were obtained from the curves compiled by the AEC Neutron Cross Section Advisory Group[5] which are based on experimental determinations; and the necessary cross sections and angular distributions for oxygen were taken from the tables prepared by Goldstein[6] which are

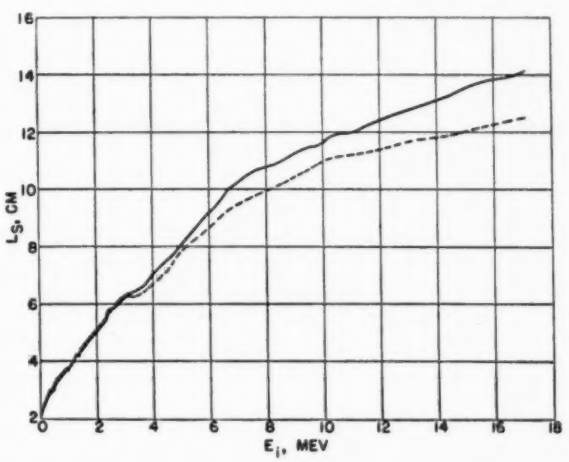


FIG. 1. Slowing-down length of neutrons in water to 1.44 ev as a function of the source energy. The solid curve is obtained when the asymmetrical nature of the O scattering is taken into account; the dashed curve is obtained when the O scattering is assumed symmetrical.

based partly on experimental results and partly on computations. The value of L_s for $E_f=1.44$ ev is shown as a function of the source energy E_i in Fig. 1 (solid curve). The same quantity as calculated by neglecting the asymmetry (i.e., setting $b_1=0$) is also shown in Fig. 1 (dashed curve). It may be seen that L_s is usually smaller in this case and that, in general, the difference between the curves increases with increasing energies, being quite small for source energies below 3.0 Mev. If L_s^2 is averaged for the thermal fission spectrum,[7] the result obtained is 27.0 cm² when the asymmetrical oxygen scattering is taken into account, which is about 3 percent higher than the value obtained if it is neglected. The effect of capture which tends to decrease L_s is included in both curves so that the curve following from Fermi's expression would lie somewhat above the dashed curve. Consequently, when the effects of asymmetry and capture are both included the calculated average value of L_s^2 exceeds that obtained from Fermi's expression by less than 3 percent. In any case the effect of asymmetry can account for only a small part of the 20 percent discrepancy between the experimental and calculated values.

[1] Hill, Roberts, and Fitch, U. S. Atomic Energy Commission Publication, AECD-3392.
[2] E. Fermi, Ricerca sci. **7**, 13 (1936).
[3] R. E. Marshak, Revs. Modern Phys. **19**, 185 (1947).
[4] H. C. Volkin, Phys. Rev. **91**, 425 (1953).
[5] U. S. Atomic Energy Commission Publication, "Neutron Cross Sections," AECU-2040.
[6] H. Goldstein, Nuclear Development Associates, Inc., unpublished Memo 15C-15.
[7] B. E. Watt, Phys. Rev. **87**, 1037 (1952).

Extension of Nonreciprocal Ferrite Devices to the 500–3000 Megacycle Frequency Range*

R. H. Fox
Lincoln Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received September 23, 1954)

ONE of the principal problems involved in the application of ferrite materials to microwave circuits has been that of extending the lower frequency limit. Recent developments in materials[1] have reduced this limit to about 3000 Mc. However, little further gain along these lines appears possible since at lower frequencies the residual magnetic loss from the ferromagnetic resonance line becomes relatively larger.

Because of this situation, the possibility of using static magnetic fields greater than that required for ferromagnetic resonance has been investigated quantitatively. Although nonreciprocal effects do diminish as the frequency decreases, the situation appears reasonably satisfactory near and above 500 Mc since component size can be allowed to increase in proportion to the wavelength.

First consider an infinite medium of ferrite magnetized in the direction of propagation. Let $H=1000$ oersteds, $4\pi M=3000$ e.m.u., dielectric constant $\epsilon=16$, and frequency $\nu \ll \nu_{\text{resonance}}=\gamma H$ $\times(\gamma=2.8$ Mc/oersted). Then, using the relations derived by Polder,[2] we find for the Faraday rotation,

$$\Phi \cong \pi \sqrt{\epsilon}\frac{4\pi M}{(H(H+4\pi M))^{\frac{1}{2}}}\left(\frac{\nu}{\nu_\gamma}\right)=\frac{\nu}{150}\frac{\text{radians}}{\text{wavelength}}.$$

Since 45° rotation is required for a circulator, this equation indicates that for $\nu>250$ Mcps devices can be made which are of the order of a wavelength long. If the field H can be further reduced without encountering excessive resonance losses, the Faraday rotation is correspondingly increased.

A calculation may also be made of the nonreciprocal phase shift in rectangular wave guide loaded with ferrite using the exact theory.[3] For this purpose it was necessary to utilize Whirlwind I Digital Computer.† Computations were made at the frequencies 3000, 1500, 750, 375 Mcps and several values of magnetic field. All dimensions were scaled as the wavelength, the equivalent slab thickness at 10 000 Mcps being 0.040 in. The length L required to produce a nonreciprocal phase shift of 90° was then calculated since this suffices for making a simple circulator. The results are shown in Table I.

TABLE I. Length of circulator as a function of frequency.

Frequency (Mcps)	Field (oersteds)	$L_{90}°$ (cm)
3000	3000	58
	2000	29
	1500[a]	18
1500	3000	143
	2000	...
	1500[a]	...
750	2000	280
	1500	185
	1000[a]	121
375	1500	782
	1000[a]	460
	500	225

[a] Expected minimum value of field giving sufficiently low (<2 percent) resonance losses.

The values of saturation magnetization and dielectric constant used in the calculations were $4\pi M=3000$ e.m.u. and $\epsilon=12.8$, respectively. By increasing the slab thickness and magnetization, it is expected that a reduction of about a factor of 3 in L can be obtained.

This mode of operation places somewhat different requirements on the ferrite materials. A narrow resonance line, high magnetization, and low dielectric loss are the principal requirements. Since low field losses cannot occur, it should be possible to tolerate moderate porosity, inhomogeneities in structure, and such sample defects.

Although it is possible that the dielectric loss problem may be a severe one at these frequencies, preliminary measurements at 425 Mc have been rather promising and will be continued.

I should like to extend my thanks to Professor C. L. Hogan of Harvard University and Dr. Benjamin Lax of Lincoln Laboratory for helpful discussions of this work and to Mr. K. J. Button of Lincoln Laboratory for programming the computer calculations.

* The research reported in this document was supported jointly by the Department of the Army, the Department of the Navy, and the Department of the Air Force under Air Force Contract No. AF 19 (122)-458.
† Digital Computer Laboratory time was made available by the Office of Naval Research.
[1] Van Uitert, Schafer, and Hogan, J. Appl. Phys. **25**, 925 (1954).
[2] D. Polder, Phil. Mag. **40**, 99 (1949).
[3] Lax, Button, and Roth, Lincoln Laboratory Tech. Mem. No. 49 (1953).

Small-Angle Scattering and Surface Structure*

YIN-YUAN LI AND R. SMOLUCHOWSKI
Carnegie Institute of Technology, Pittsburgh, Pennsylvania
(Received September 18, 1954)

SINCE submicroscopic irregularities on the surface of a solid can be regarded as fine particles attached to a plane substrate, it follows that they may cause diffuse scattering of x-rays (or thermal neutrons) in the same way as powders. In principle, this effect would cause a broadening of all the Bragg diffraction maxima, but in practice its intensity may become observable only in the case of small-angle scattering. In the following we shall consider the latter case.

For a mathematical analysis and a numerical estimate, let us take a flat specimen whose volume V extends from $-A/2$ to $A/2$, from $-B/2$ to $B/2$, and from $-C/2-\delta_2(x,y)$ to $C/2+\delta_1(x,y)$ along the x, y, and z axis, respectively, where A, B, and C are constants, and δ_1 and δ_2 are small and positive. Let I_e be the intensity of the radiation scattered by a single electron and n the number of electrons per unit volume in the specimen. The amplitude of the scattered radiation is then given by

$$A=(I_e)^{\frac{1}{2}}n\int\int\int_V \exp[i2\pi(k_x x+k_y y+k_z z)]dxdydz, \quad (1)$$

where $\mathbf{k}=(\mathbf{s}-\mathbf{s}_0)/\lambda$, $\mathbf{s}$ and $\mathbf{s}_0$ are unit vectors in the directions of

the scattered and incident rays, respectively, and λ is the wavelength of the radiation. Making the approximation $\exp i2\pi k_z = 1+2\pi i k_z$ we have

$$A=(I_e)^{\frac{1}{2}}n\Big\{ABCf(k_xA)f(k_yB)f(k_zC)+\exp[i\pi k_zC]\times\int\int\delta_1(x,y)\exp[i2\pi(k_xx+k_yy)]dxdy+\exp[-i\pi k_zC]\times\int\int\delta_2(x,y)\exp[i2\pi(k_xx+k_yy)]dxdy\Big\},\quad(2)$$

where $f(u)$ stands for $(\sin\pi u)/\pi u$. Let us take the idealized case where all the humps are of rectangular shape with a height c_i and extend from $A_i-a_i/2$ to $A_i+a_i/2$ and from $B_i-b_i/2$ to $B_i+b_i/2$ on the upper surface and similarly those with index j on the lower surface. We have then

$$A=(I_e)^{\frac{1}{2}}n\{ABCf(k_xA)f(k_yB)f(k_zC)+\Sigma_i a_ib_ic_i\times\exp[i2\pi(k_xA_i+k_yB_i+k_zC/2)]f(k_xa_i)f(k_yb_i)+\Sigma_j a_jb_jc_j\times\exp[i2\pi(k_xA_j+k_yB_j-k_zC/2)]f(k_xa_j)f(k_jb_j)\}.\quad(3)$$

First we want to show that in the experimental range of the scattering angle the contribution from the surface I_S can be more important than that from the body I_B and second that its intensity is observable when a bent crystal monochromator and a counter recording technique is used. In order to do so we need to discuss the following intensity terms:

$$I_B/I_en^{-2}=[ABCf(k_xA)f(k_yB)f(k_zC)]^2,\quad(4)$$

and

$$I_S/I_en^{-2}\sim N[abcf(k_xa)f(k_yb)]^2,\quad(5)$$

where N is the total number of humps on the irradiated surface and a, b, and c are the average values of a_i, b_i, and c_i, respectively. In writing Eq. (5) we assume that the surface irregularities are sparse and are located at random in such a manner that we may neglect the interference between the waves from different humps. The following numerical example corresponds to typical experimental conditions in some of our work: The incident monochromatic beam of x-rays is in the direction of the z axis, and measurements are taken in the region of the scattering angle $\epsilon=10^{-2}$ radian. Further, $\lambda=1.5\times10^{-8}$ cm, $A=B=3\times10^{-1}$ cm, $C=10^{-2}$ cm, and $a=b=c=10^{-5}$ cm. We have, then, $k_x\sim k_y\sim\epsilon/\lambda=\frac{2}{3}\times10^6$, $k_z\sim\epsilon^2/2\lambda=\frac{1}{3}\times10^4$, and

$$I_B/I_en^2=1.4\times10^{-30},$$
$$I_S/I_en^2=\ 5\times10^{-34}N.$$

Making the assumption that one percent of the surface area is covered by irregularities, we have

$$N=AB/ab\times1/100\sim10^7,$$

and so $I_S=10^3I_B$. If this is the case, we may also neglect the cross term between the scattered amplitudes from the main body and from the surface, i.e. we may simply omit the first term in Eq. (3). With a bent quartz monochromator and a conventional x-ray tube one can obtain $I_0=10^9$ cm^{-2} sec^{-1} photons at the sample position. At the counter position 20 cm away $I_e=2\times10^{-10}$ cm^{-2} sec^{-1} photons, and so $I_S=10^3$ cm^{-2} sec^{-1} photons, assuming $n=10^{24}$ cm^{-3}. With a counter slit 0.005 in. wide and 0.5 in. high one obtains I_S of the order of 15 counts per second, which can be well observed above the usual cosmic-ray background. It should be noted that this scattering is not altered by usual polishing agents since their grain is too coarse to affect the small surface irregularities here considered.

The small-angle scattering offers a possible method for measuring the average size of surface irregularities. This could be of value when the application of electron microscopy or of optical interferometry is not advisable, for instance when the presence of a replica or the coating by a reflecting film disturbs the surface. As an example the method may be useful in a study of nucleation on surfaces. The quantitative analysis of experimental data on surface scattering is essentially the same as the analysis of scattering by particles.[1] For instance, for surface scattering Guinier's approximate formula should read

$$I(\epsilon)=I_eN\eta^2\exp[-2\pi^2\epsilon^2R^2/\lambda^2],$$

where N is the total number of surface humps under irradiation, η the average number of electrons in one of these humps, and

$$R^2=\frac{n}{\eta}\int_A\delta(x,y)(x^2+y^2)dxdy.$$

The integration is taken parallel to the surface and over the cross section area A of a typical surface hump. The origin of the coordinates is now at center of gravity of the hump and $\delta(x,y)$ is its height. R represents the average radius of the cross section A and can be obtained by plotting the experimental $\ln I$ *vs* ϵ^2. The average value of the height $\delta\sim\eta/nR^2$ cannot be unambiguously determined, since the experimental data give $N\eta^2$ rather than η.

The presence of surface scattering brings complications into the interpretation of the small-angle scattering data in a study of the lattice imperfections of solids. For example, results obtained on deformed copper indicate the presence of scattering centers of[2] linear dimensions of 10^2–10^3 A. It is rather difficult to understand what internal lattice imperfections in a pure metal could be so large. On the other hand, the occurrence of this size surface irregularities or precipitates, if the metal were impure, is quite reasonable. Furthermore, if the scattering is caused by surface defects, then a surface diffusion process may cause a rapid variation of size of these defects upon annealing. In any case one must be careful in interpreting small-angle scattering as caused only by the internal lattice imperfections. This is especially true when the scattered intensity is weak making the surface contribution not negligible.

* Research supported by U. S. Atomic Energy Commission contract.

[1] For a brief review of the small-angle x-ray scattering see Chapter 12 of A. Guinier's *Radiocrystallographie* (Dunod, Paris, 1945). An extensive compilation of references on this subject is given by K. L. Yudowitch in *Bibliography of Small-angle X-ray Scattering* (American Crystallographic Association, 1952).

[2] S. Hayes and R. Smoluchowski, Appl. Sci. Research (B)4, 10 (1954).

Analog Methods for Study of Transient Heat Flow in Solids with Temperature-Dependent Thermal Properties*

NORMAN E. FRIEDMANN
Department of Engineering, University of California, Los Angeles, California

(Received June 21, 1954)

THE purpose of this note is to describe methods in which readily available nonlinear semiconductors[1] are used as circuit elements in analog computers of the electronic[2] or circuit[3] type to obtain the solution of nonlinear heat transfer problems. The proposed methods differ from those existing in that they do not require prohibitive quantities of costly function generators and multiplying equipment[4,5] or mechanically variable circuit elements.[6]

If we approximate the spatial derivative by a divided central difference[7] in the equation for unidimensional heat flow in an isotropic stationary medium with no internal heat sources and surface losses, we obtain

$$\phi_{n-1}-2\phi_n+\phi_{n+1}=\langle\Delta x\rangle^2\frac{d}{d\theta}\left[\left\{\frac{1}{\phi_n}\int_0^{\phi_n}\frac{d\lambda}{\alpha(\lambda)}\right\}\cdot\phi_n\right].\quad(1)$$

In this expression ϕ is a transformed temperature related to the actual temperature T by

$$\phi=\int_0^T K(\lambda)d\lambda.\quad(2)$$

K is thermal conductivity, $\alpha=K/\rho C_P$ is thermal duffusivity, θ is time, x is spatial position, and λ is a dummy variable.

In the electronic analog computer the nonlinearity appearing in

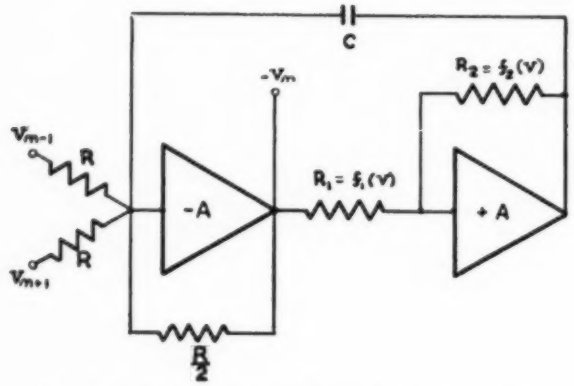


FIG. 1. Electronic computer cell for nonlinear heat transfer.

Eq. (1) is accounted for by a nonlinear resistance. The measured voltage change with time at each node can then be directly correlated with temperature through Eq. (2). The required resistance variation can be determined by writing the equation for the single cell shown in Fig. 1.† This cell consists of two standard operational amplifiers,[8] one of negative gain and the other of positive gain, and two semiconductor networks R_1 and R_2.‡ We can write, neglecting the input to the amplifiers,

$$V_{n-1}-2V_n+V_{n+1}=RCd/dt(R_2/R_1V_n), \tag{3}$$

which, when compared with Eq. (1), results in the equation of condition on R_1 and R_2

$$\frac{R_2}{R_1}\approx\frac{1}{\phi_n}\int_0^{\phi_n}\frac{d\lambda}{\alpha(\lambda)}. \tag{4}$$

The voltages VR_1 and VR_2 are given by

$$VR_1=-V_n \quad VR_2=-\frac{R_2}{R_1}V_{n_0}. \tag{5}$$

In the electric circuit type computer the nonlinear term appearing in Eq. (1) can be accounted for by a nonlinear capacitance. As previously, measured voltage dependency on time can be correlated with temperature. The required capacitance variation can be found by writing the difference equation for node voltages for the single cell shown in Fig. 2, *viz.*,

$$V_{n-1}-2V_n+V_{n+1}=Rd/dt(C_nV_n), \tag{6}$$

from which it follows by comparison with Eq. (1) that C_n must satisfy the condition

$$C_n\approx\frac{1}{\phi_n}\int_0^{\phi_n}\frac{d\lambda}{\alpha(\lambda)}.$$

A circuit arrangement which makes use of function generators to obtain a capacitance which is a function of the voltage difference across its terminals is described in reference.[9] A slightly different circuit for obtaining a non-linear capacitance can be obtained as a modification of that shown in Fig. 1. The resultant equations are similar to Eqs. (4) and (5).

To evaluate the characteristics of the electronic computer method the integrated diffusivity dependence on temperature of 302 stainless steel (Eq. (1)) was computed from available data.[10,11] The change of R_2 with $V_n(\phi_n)$ called for was then found by means of Eq. (4). These results served as a basis for design of the semiconductor network. The experimental calibration of the

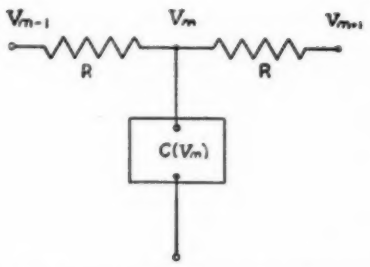


FIG. 2. Electric computer cell for nonlinear heat transfer.

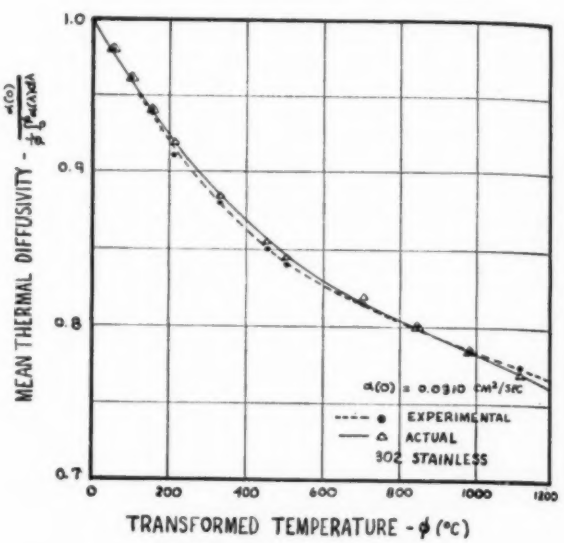


FIG. 3. Thermal diffusivity of 302 stainless steel.

network is compared in Fig. 3 with the characteristics required. The reproduction is seen to be within $\pm\frac{1}{2}$ percent.

Work is in progress to solve a number of representative problems by both methods. In each case a number of the basic cells will be connected to minimize lumping errors[12] which arise as a result of approximation of the derivative $\partial^2\phi/\partial x^2$ by finite differences.

* Work sponsored by University Research Grant.

† This is a modification of a method suggested to the author by Dr. M. Tribus, University of California, Los Angeles, Department of Engineering.

‡ These are characterized by a positive resistance which monotonically decreases with increasing voltage or current, such as germanium or silicon carbide.

[1] K. B. McEachron, Gen. Elec. Rev., February, 1930, pp. 92–99.
[2] A. B. MacNee, Proc. Inst. Radio Engrs. **37**, 1315–1324 (1949).
[3] McCann, Wilts, and Locanthi, Proc. Inst. Radio Engrs. **37**, 954–961 (1949).
[4] C. A. Meneley, and C. D. Morrill, Proc. Inst. Radio Engrs., Computer Issue, **41**, 1487–1496 (1953).
[5] G. D. McCann, and C. H. Wilts, J. Appl. Mech. **16**, 247–258 (1949).
[6] V. Paschkis, Trans. Am. Inst. Elec. Engrs. **68**, 70–73 (1949).
[7] R. M. Howe and V. S. Haneman, Proc. Inst. Radio Engrs., Computer Issue, **41**, 1497–1508 (1953).
[8] G. A. Korn and T. M. Korn, *Electronic Analog Computers* (McGraw-Hill Book Company, Inc., New York, 1952).
[9] McCann, Wilts, and Locanthi, Trans. Am. Inst. Elec. Engrs. **68**, 652–660 (1949).
[10] *Metals Handbook* (Am. Soc. Metals, Cleveland, Ohio, 1949).
[11] L. Silverman, J. Metals **5**, 631–632 (1953).
[12] Klein, Touloukian, and Eaton, Am. Soc. Mech. Engrs. Paper No. 52-A-65, presented at Annual Meeting, New York, New York, Nov. 30–Dec. 5, 1952.

Books Reviewed

Proceedings for Short Courses on High Temperature Properties of Materials and Mechanics of Creep

Proceedings are now available on short courses held at The Pennsylvania State University during June, 1954, dealing with high-temperature properties of materials and mechanics of creep.

The proceedings on the "High Temperature Properties of Materials" course were prepared by Dr. C. Zener, Dr. B. J. Lazan, Dr. G. V. Smith, Mr. M. J. Manjoine, Mr. W. D. Manly, and Mr. D. R. Miller.

The proceedings on the "Mechanics of Creep" short course were prepared by Dr. J. E. Dorn, Dr. G. R. Irwin, Dr. L. F. Coffin, Professor F. R. Shanley, Mr. G. D. Lubahn, and Mr. E. A. Davis.

For further information regarding these proceedings, write Dr. Joseph Marin, Department of Engineering Mechanics, The Pennsylvania State University, State College, Pennsylvania.

Introduction to Atomic and Nuclear Physics

By Henry Semat, Pp. 561+xii, 3rd Edition. Rinehart & Company, Inc., New York. Price $6.50.

The principal change in this third edition of *Introduction to Atomic Physics* is revealed by the title. Section 3, on nuclear physics, has been doubled in length, with the addition of chapters on fundamental particles (especially mesons and V particles), new elements and isotopes, and particle accelerators. The previous chapters on natural radioactivity, disintegration of nuclei, nuclear processes, and nuclear fission have been enlarged and brought up to date.

The other two sections of the book, "Foundations of Atomic and Nuclear Physics," and "The Extranuclear Structure of the Atom," have been modified by the inclusion of some new material and the creation of three new chapters by reorganization. Beginning with a review of the principles of electricity and magnetism, Section 1 proceeds to discuss charged particles, radiation, and wave-particle dualism. Section 2 follows with a more detailed description of the hydrogen atom, and the electron distribution of heavier atoms and its relation to optical and x-ray spectra.

At the back of the book are located seven appendixes, including an up-to-date table of isotopic masses.

Introduction to Atomic and Nuclear Physics is designed as a text for an undergraduate course in modern physics for students with at least one year of elementary physics and some calculus.

Analysis of Deformation. Volume I. Mathematical Theory

By K. Swainger. Pp. 285+xix. The Macmillan Company, New York, 1954. Price $12.75.

The primary purpose of this volume, the first of two on the subject of deformation, is to ". . . formulate the mathematically linear theory analysing deformation and to draw general inferences without solving the equations and boundary conditions for particular cases." The application and experimental examination of the theory constitute the subject matter of the second volume.

In discussing the problem of deformation, the author has not considered the atomistic approach but, rather, treats a substance as a continuum. He has attempted to apply a comprehensive, linear, phenomenological theory to the analysis of finite deformation of all solids. The theory utilizes three reference frames, one of which is deformable, to describe the displacement of a physical point. It is shown that the rotation effects caused by the strain field need not be dealt with.

Some of the broader topics discussed are displacement, stress, strain, thermal considerations, isotropic elastic substances, yield, and viscoelastic and elastoplastic isotropic substances. A treatment of vector analysis (used extensively throughout the book) occupies about thirty pages of the appendix.

This volume may conceivably serve both as text and reference work at a somewhat advanced level.

Modern Physics for the Engineer

Edited by Louis N. Ridenour. Pp. 499+xix. McGraw-Hill Book Company, Inc., New York, 1954. Price $7.50.

This volume contains a series of lectures by eighteen prominent scientists given at the University of California at Los Angeles in the year 1952–1953. The purpose of these lectures was to familiarize engineering students and practicing engineers with some of the most recent developments in physics. The articles in this book are descriptive, assuming only a familiarity with the terminology of physics.

After an introduction by R. Weller, the book is divided into three parts. The first part restricts itself to fundamental physics and includes articles by H. P. Robertson on "Relativity and Mechanics," L. I. Schiff on "Atomic Structure," F. Seitz on "Physics of the Solid State," C. Kittel on "Magnetism," W. D. Hershberger on "Microwave Spectroscopy," W. A. Fowler on "Nuclear Structure and Transmutation," R. V. Langmuir on "Electronuclear Machines," G. T. Seaborg on "The Actinide Elements and Nuclear Power," and W. K. H. Panofsky on "Elementary Particles."

The second part of this book concerns itself with the application of physics to our natural environment. It contains articles by J. L. Greenstein on "Astrophysics," D. T. Griggs on "High Pressure Phenomena with Applications to Geophysics," R. Revelle on "The Earth beneath the Sea," L. B. Loeb on "Thunderstorms and Lightning Strokes," and W. Bleakney on "Transient Phenomena in Supersonic Flow."

The third part, which covers the application of physics to information and communication, includes articles by S. Ramo on "Electrons and Waves," J. Bardeen on "Semiconductor Electronics," J. B. Wiesner on "Communication Theory and the Transmission of Information," and L. N. Ridenour on "Computing Machines and the Processing of Information."

Table of Sine and Cosine Integrals for Arguments from 10 to 100

National Bureau of Standards Applied Mathematics Series 32 (reissue of Mathematical Tables 13). Pp. 186. Price $2.25. (Order from the Government Printing Office, Washington 25, D. C.).

The sine and cosine integrals have long played an important part in the theory of numbers and the calculus of probabilities. Lately, they have assumed increasing value in such fields as antenna theory, electromagnetic theory, and nuclear physics.

This table contains $\mathrm{Si}(x)$ and $\mathrm{Ci}(x)$, $x=10(0.01)100, 10D$. In addition there are auxiliary tables of multiples of $\pi/2$, $(\frac{1}{2}p)(1-p)$, $p(1-p^2)/6$, and $q(1-q^2)/6$, where $p+q=1$, to facilitate interpolation in the tables.

Electronic Measuring Devices

By E. H. W. Banner, Pp. 395+xiv. Chapman and Hall Ltd., London, 1954. Price 45 shillings.

This volume contains a survey of the field of instruments using electronics to measure electrical as well as nonelectrical quantities. Brief descriptions of many types of instruments are presented. Examples of specific instruments, chosen almost entirely from those manufactured in England, are described. This book is designed for use by instrument engineers, and for those who wish to know the principles behind their instruments. It will be of interest to students with some knowledge of electronics, but it is not a textbook.

The first part of this book describes the indicating instruments by which the final results are presented, such as galvanometers and recording devices. The second part introduces the principles

of electronic tubes and the electronic circuits used in instruments. It also includes descriptions of the cathode-ray tube, photoelectric and semiconductor devices, and radiation detectors.

The third and fourth sections show how the devices described in the first two parts can be used for specific types of measurements. A very wide variety of instruments and measuring techniques are described. For example, radiation measuring devices, strain gauges, proximity meters, vacuum gauges, *p*H meters, and photometers, as well as many varieties of instruments for measuring electrical quantities, are discussed. Each chapter contains a bibliography for further reference.

Quantum Mechanics

By F. Mandl. Pp. 233+viii. Academic Press, Inc., New York, 1954. Price $5.80.

Designed for both the student and the research worker, this book presents the basic elements of quantum mechanics without stressing mathematical rigor. With a geometrical formulation in mind, the author develops the mathematical formalism of quantum mechanics in the first five chapters. After a brief mathematical introduction in Chapter I the topics of wave-mechanical concepts, energy eigenfunctions, general quantum mechanical principles, and matrix mechanics are treated in turn.

The remaining four chapters are devoted to illustrations of how the general theory may be applied to several outstanding problems. Systems of many particles, with emphasis on angular momentum, spin, and symmetry properties, are treated in great detail. Time-independent perturbation theory and collision processes are discussed along with the basic ideas of group theory.

Excercises are given at the end of each chapter and a section providing detailed hints for their solution is included.

A knowledge of classical physics, calculus, and vector analysis as well as of the qualitative concepts of quantum theory is presupposed.

A Treatise on Electricity and Magnetism

By J. C. Maxwell. Volume One: Pp. 519+xxxii. Volume Two: Pp. 507+xxiv. Dover Publications, New York, 1954. Price $4.95.

These two volumes, bound as one, are an unabridged republication of the third edition of 1891 of James Clerk Maxwell's work. Reproduced, in unaltered form, are all four parts of the *Treatise* —electrostatics, electrokinematics, magnetism, and electromagnetism—complete with 107 figures and 20 full-page plates.

It is the expressed intention of the publisher to promote as widespread a circulation of the book as possible, especially among teachers and students. It is for this reason, he states, that the book is being sold at the price quoted.

Thermal Conductivity of Metals and Alloys at Low Temperatures

By R. L. Powell and W. A. Blanpied. Pp. 68. National Bureau of Standards Circular No. 556. Price $.50. (Order from Government Printing Office, Washington 25, D. C. Cost outside of U. S., $.65 in U. S. exchange).

This circular was prepared by the NBS-AEC Cryogenic Laboratory at Boulder. It includes data on construction materials for cryogenic apparatus. The more important data are plotted in 48 graphs. There is some data on dielectrics.

Permissible Dose from External Sources of Ionizing Radiation

National Bureau of Standards Handbook 59. Pp. 79. Price $.30. (Order from the Government Printing Office, Washington 25, D. C.).

The recommendations and discussions of permissible dose contained in this handbook form the basis of all other recommendations of the National Committee on Radiation Protection. The handbook presents permissible doses for radioactive material within the body, and information on the safe handling of radioactive materials and waste disposal. Although the formal preparation of this report has been delayed, the substance of the recommendations of the Subcommittee on Permissible Dose from External Sources has been well known and in use for a number of years.

The handbook presents discussions of the basic concepts of permissible dose and discussions of each of the many factors considered in the formulation of the recommendations. For easy reference the exposure limits of parts of the body to various types of ionizing radiation are briefly stated in the section "Protection Rules."

Theoretical Elasticity

By A. E. Green and W. Zerna. Pp. 442+xiii. Oxford University Press, New York, 1954. Price $8.00.

Three aspects of elasticity theory which have received considerable attention recently are treated at an advanced level. The main topics are: (1) finite elastic deformations, (2) complex variable methods for two-dimensional problems for isotropic and for nonisotropic bodies, and (3) shell theory. The discussion of the last two topics is limited to the domain of classical infinitesimal elasticity. General tensor notation is used throughout, enabling the general theories to be expressed in an elegant and compact form.

The first chapter of mathematical preliminaries contains a summary of tensor analysis restricted to a three-dimensional Euclidean space. The general theory of elasticity for finite deformations is treated next, with special attention being given to the stress-strain relations for an isotropic body. Solutions of a number of special problems, mostly for incompressible isotropic bodies, are obtained, usually in a form that is independent of the strain-energy function. Small deformations superposed on finite deformations are considered and a theory treating both compressible and incompressible bodies is presented.

The classical infinitesimal theory is deduced as a special case of the general theory with stress-strain relations for both isotropic and nonisotropic bodies. Plane strain and plate theories are first developed in tensor notation. These are then specialized and the complex variable notation introduced. The powerful techniques of complex function theory are then utilized in solving the two-dimensional problems.

The last five chapters are devoted to shell theory. Emphasis is placed on formulation of a general theory.

Atomic Energy and Its Applications

By J. M. A. Lenihan. Pp. 265+xiii. Pitman Publishing Corporation, London, 1954. Price $4.00.

This book was designed ". . . to provide a simple account of the physical foundations of nuclear science, a concise description of their applications, and a guide to future progress in the exploitation of atomic energy." As such, it is a book for those not versed in the complexities of nuclear physics. It is intended for those with an undergraduate understanding of the principles of physics. The rather broad scope of the work implies an attempt on the part of the author to convey much information in a relatively small amount of space to those who desire a reference book on a branch of science in which they are not personally engaged.

The book is divided, essentially, into two sections of roughly equal length. The first presents the underlying tenets of atomic and nuclear physics. Included are discussions of atomic structure, waves and atoms, radioactivity, nuclear structure, nuclear reactions, and fission. In addition there is a chapter dealing with the detection and measurement of nuclear radiations and one on particle accelerators.

The second half of the book is devoted to applications. The topics treated here are reactors, atomic weapons, production of

nuclides, radiation hazards, medical, industrial, and scientific applications of radioactive materials, and future prospects of atomic energy.

Tensor Analysis for Physicists

By J. A. SCHOUTEN. Pp. 277+xii. 2nd Edition. Oxford University Press, New York, 1954. Price $4.80.

This text is based on lectures given by the author at Amsterdam University and is divided into two main sections, theory and applications. Chapters I and II deal with tensor algebra in E_n and R_n, the affine space and the space associated with all orthogonal homogeneous transformations, respectively. The relations between matrix calculus and tensor calculus are also developed in Chapter II. Chapters IV and V are devoted to tensor analysis in X_n, the manifold in which all invertible analytic transformations are allowed, and in L_n, which is an X_n with a linear displacement. The identification of quantities in E_n is contained in Chapter III.

A summary of Chapters I–V is included for the benefit of those who are mainly concerned with applications dealt with in the latter part of the book and who do not wish to go through the mathematical detail of the earlier chapters.

The applications treated include the theory of elasticity, classical dynamics, and relativity. Dirac's matrix calculus is discussed in the last chapter.

There are exercises at the end of each chapter, many of which are formulated as proofs so as to provide an answer in advance.

Other Books Received

Die Grundlagen der Akustik. EUGEN SKUDRZYK. Pp. 1084 +xxii, Figs. 450. Springer-Verlag, Berlin, Wien 1, 1954. Price $35.00.

Niederdruck-Stromrichterventile. HANS VON BERTELE. Pp. 239+xii, Figs. 149. Springer-Verlag, Berlin, Wien 1, 1953. Price $9.30

Grundlagen der Elektronenoptik. WALTER GLASER. Pp. 699+x, Figs. 445. Springer-Verlag, Berlin, Wien 1, 1953. Price $28.60.

Kurzes Lehrbuch der Elektrotechnik. GUNTHER OBERDORFER. Pp. 199+vii, Figs. 231. Springer-Verlag, Berlin, Wien 1, 1952. Price $3.50.

Fernsehtechnik. F. KIRSCHSTEIN AND G. KRAWINKEL. Pp. 288, Figs. 231, 5 Plates. S. Hirzel Verlag, Stuttgart, 1952.

Advanced Theory of Waveguides. L. LEWIN. Pp. 192. Iliffe and Sons, Dorset House, Stamford Street, London, 1952.

Direct Current Machines for Control Systems. A. TUSTIN. Pp. 306+xii, with Figs. The Macmillan Company, New York, 1952. Price $10.00.

Advances in Geophysics, Volume 1. Edited by H. E. LANDSBERG. Pp. 362+xi, Figs. 18. Academic Press, New York, 1952. Price $7.80.

Photoconductivity in the Elements. TREVOR S. MOSS. Pp. 263+x, with Figs. Academic Press, New York, 1952. Price $7.00.

Optics, Vol. IV. ARNOLD SOMMERFELD. Pp. 383+xiii, Figs. 97. Academic Press, New York, 1954. Price $6.80.

Communication Theory. Edited by WILLIS JACKSON. Pp. 532+xii, with Figs. Academic Press, New York, 1953. Price $11.00.

Television and Radar Encyclopedia. Edited by W. MACLANACHAN. Pp. 216+viii, Figs. 225. Pitman Publishing Corporation, New York, 1953. Price $6.00.

Electron Optics. O. KLEMPERER. Pp. 471+xii, with Figs. Cambridge University Press, New York, 1953. Price $9.50.

Electronics. ARTHUR O. WILLIAMS. Pp. 306+v, D. Van Nostrand Company, New York, 1953. Price $4.50.

Flow Properties of Disperse Systems. J. J. HERMANS. Pp. 445+xi, with Figs. Interscience Publishers, Inc., New York, London, 1954. Price $9.90.

Psychology of Invention in the Mathematical Field. JACQUES HADAMARD. Pp. 145+xiii, Dover Publications, New York, 1954. Price $1.25; Clothbound $2.50.

Microphysical World. WILLIAM WILSON. Pp. 216+vii, with Figs. Philosophical Library, New York, 1954. Price $3.75.

Modern Science and God. P. J. MCLAUGHLIN. Pp. 89. Philosophical Library, New York, 1954. Price $2.75.

Development of the Guided Missile. KENNETH W. GATLAND. Pp. 133+x, Figs. 45. Philosophical Library, New York, 1952. Price $3.75.

Electronic Measuring Instruments. E. H. W. BANNER. Pp. 395+xiv, with Figs. Chapman and Hall Ltd., London W.C. 2, 1954.

Statistical Methods in Electrical Engineering. D. A. BELL. Pp. 175+viii, with Figs. Chapman and Hall, Ltd., London W.C.2, 1953.

Calculation of Compressible Airflow. L. ROSENHEAD. Pp. 115+x, Graphs, Oxford University Press, New York, 1954. Price $13.45.

Theoretical Elasticity. A. E. GREEN AND W. ZERNA. Pp. 442+xiii, with Figs. Oxford University Press, New York, 1954. Price $8.00.

Hardness of Metals. D. TABOR. Pp. 175+ix, Figs. 63, Oxford University Press, New York, 1953. Price $3.00.

Stress Waves in Solids. H. KOLSKY. Pp. 211+x, Figs. 46, 3 plates, Oxford University Press, New York, 1953. Price $5.00.

Micrometeorology. O. C. SUTTON. Pp. 333+xii, Figs. 35. McGraw-Hill Book Company, New York, 1953. Price $8.50.

Fluid Dynamics, Volume IV. M. H. MARTIN. Pp. 186+v. McGraw-Hill Book Company, New York, 1953. Price $7.00.

Ionic Processes in Solution. RONALD W. GURNEY. Pp. 275 +ix, Figs. 77. McGraw-Hill Book Company, New York, 1953. Price $6.50.

Introduction to Geometrical and Physical Optics. JOSEPH MORGAN. Pp. 450+xi, with Figs., McGraw-Hill Book Company, New York, 1953. Price $6.50.

Engineering Statistics and Quality Control. IRVING W. BURR. Pp. 442+xi, with Figs. McGraw-Hill Book Company, New York, 1953. Price $7.00.

Elementary Differential Equations. LYMAN W. KELLS. Pp. 266+x, with Figs. McGraw-Hill Book Company, New York, 1953. Price $4.00.

Static and Strength of Materials. ROLAND H. TRATHEN. Pp. 506+x, with Figs. John Wiley and Sons, Inc., New York, 1954. Price $7.50.

Metadyne Statics. JOSEPH M. PESTARINI. Pp. 318, John Wiley and Sons, Inc., New York, 1953. Price $9.00.

Man and His Physical Universe. RICHARD WISTAR. Pp. 303, 9 Tables. John Wiley and Sons, Inc., New York, 1953. Price $4.75.

Physical Formulae. T. S. E. THOMAS. Pp. 118+vii, Figs. 7. John Wiley and Sons, Inc., New York, 1953. Price $2.00.

Microwave Spectroscopy. W. GORDY, W. V. SMITH, AND R. F. TRAMBARULO. Pp. 446+xii, Figs. 95. John Wiley and Sons, Inc., New York, 1953. Price $8.00.

Elements of Heat Treatment. G. M. ENOS AND W. E. FONTAINE. Pp. 286, with Figs. John Wiley and Sons, Inc., New York, 1953. Price $5.00.

Introduction to Aeronautical Dynamics. MANFRED RAUSCHER. Pp. 664+xv, Figs. 479. John Wiley and Sons, Inc., New York, 1953. Price $12.00.

Introduction to Power System Analysis. FREDERICK S. ROTHE. Pp. 187+ix, Figs. 155. John Wiley and Sons, Inc., New York, 1953. Price $5.00.

Alternating Current Measurements. DAVID OWEN. Pp. 120 +vii, Figs. 80. John Wiley and Sons, Inc., New York, 1953. Price $2.00.

Microwave Lenses. J. BROWN. Pp. 125+xi, with Figs. John Wiley and Sons, Inc., New York, 1953. Price $2.00.

Essentials of Engineering Thermodynamics. HERMAN J. STOEVER. Pp. 279+vi, Figs. 72. John Wiley and Sons, Inc., New York, 1953. Price $4.50.

Procedures in Experimental Metallurgy. A. U. SEYBOLT AND J. E. BURKE. Pp. 340+xi, with Figs. John Wiley and Sons, Inc., New York, 1953. Price $7.00.

Measurement of Radio Isotopes. DENIS TAYLOR. Pp. 118+viii, with Figs. John Wiley and Sons, Inc., New York, 1952. Price $1.50.

Electrical Communication Experiments. H. R. REED, T. C. G. WAGNER, and G. F. CORCORAN. Pp. 458+viii. John Wiley and Sons, Inc., New York, 1952. Price $6.75.

Introduction to the Chemistry of the Hydrides. DALLAS T. HURD. Pp. 231+x. John Wiley and Sons, Inc., New York, 1952. Price $5.50.

Mechanics. Part 1: Statics; Part 2: Dynamics. J. L. MERRIAM. Pp. 671, John Wiley and Sons, Inc., New York, 1952. Price $4.00 for each volume.

Comets and Meteor Streams. J. G. PORTER. Pp. 123+xi. Figs. 26. John Wiley and Sons, Inc., New York, 1952. Price $5.25.

Inorganic Chemistry. THERALD MOELLER. Pp. 966+ix. John Wiley and Sons, Inc., New York, 1952. Price $10.00.

Photoelectric Tubes. A. SOMMER. Pp. 118+viii, Figs. 27. John Wiley and Sons, Inc., New York, 1952. Price $1.90.

Elasticity in Engineering. ERNEST E. SECHLER. Pp. 419+ix, with Figs. John Wiley and Sons, Inc., New York, 1952. Price $8.50.

Engineering Materials. MELVIN NORD. Pp. 518+x, Figs. 172. John Wiley and Sons, Inc., New York, 1952. Price $6.50.

Filter Design Data for Communication Engineers. J. H. MOLE. Pp. 252+xvi, Figs. 127. John Wiley and Sons, Inc., New York, 1952. Price $7.50.

Computing Methods and the Phase Problem in X-Ray Crystal Analysis. RAY PEPINSKY. Pp. 399+17, with Figs. X-Ray Crystal Analysis Laboratory, State College, Pennsylvania, 1952. Price $7.50.

Bibliography and Abstracts on Electrical Contacts. DEAN HARVEY, Honorary Chairman Committee B-4. Pp. 258+v, American Society for Testing Materials, Philadelphia, 1953. Price $5.50.

Organisation of Applied Research in Europe, the United States and Canada. Volume I, A Comparative Study. Price $1.00. Volume II, Applied Research in Europe. Price $2.00. Volume III, Applied Research in the United States and Canada. Price $1.25. Available at: The Organisation for European Economic Cooperation Mission, Publications Office, 2002 P Street, N.W., Washington 6, D. C.

General Theory of High Speed Aerodynamics. Volume VI, High Speed Aerodynamics and Jet Propulsion. W. R. SEARS. Pp. 758+xiv, with Figs. Princeton University Press, Princeton, New Jersey, 1954. Price $15.00.

Physical Measurements in Gas Dynamics and Combustion. Volume IX, High Speed Aerodynamics and Jet Propulsion. R. W. LADENBURG. Pp. 578+xvi, illus. Princeton University Press, Princeton, New Jersey, 1954. Price $12.50.

Nuclear Species. H. E. HUNTLEY. Pp. 193+xix, with Figs. St. Martin's Press, New York, 1954. Price $4.50.

Announcement

Third U. S. National Congress of Applied Mechanics Preliminary Announcement

The Third U. S. National Congress of Applied Mechanics will be held at Brown University, Providence, Rhode Island, during June 11–14, 1958. It is hoped that the scheduling of conflicting meetings can be avoided by this early announcement of the date chosen for the Congress. Further announcements concerning the preparation of papers will be made as the Congress draws nearer.

Inquiries regarding the Congress should be addressed to one of the following members of the Organizing Committee at Brown University, Providence 12, Rhode Island: Professor D. C. Drucker, Secretary; Professor E. H. Lee, Treasurer; Professor W. Prager, Chairman.